# Small Organic Molecules as Tunable Tools for Biology

**Dissertation**

**zur**

**Erlangung der naturwissenschaftlichen Doktorwürde**
**(Dr. sc. nat.)**

**vorgelegt der**

**Mathematisch-naturwissenschaftlichen Fakultät**

**der**

**Universität Zürich**

**von**

Andrea Unzue Lopez

**aus**

Spanien

**Promotionskomitee**

Prof. Dr. Cristina Nevado (Vorsitz und Leitung der Dissertation)

Prof. Dr. Amedeo Caflisch

Prof. Dr. Damian Brunner

**Zürich, 2015**

# CONTENTS

# ACKNOWLEDGMENTS

First and foremost, I would like to thank Prof. Dr. Cristina Nevado for giving me the opportunity to be part of her group. It has been an honor to have worked under her supervision. I truly appreciate her motivation and scientific commitment throughout these years and everything I have learned from her. My PhD work is a result of a fine balance between her encouraging supervision and independency to pursue my own ideas.

I would also like to thank Prof. Dr. Amedeo Caflisch for his continuous enthusiasm and inspiring discussions during our numerous collaboration projects.

This work would not have been possible without the outstanding infrastructure and service provided by the NMR and MS teams of the Department of Chemistry. In addition, I would like to thank the Center for Microscopy and Image Analysis of the UZH for providing constant help.

During my PhD studies I have been delighted to share the lab with many members of the Nevado group. I would specially like to thank Karine, David and Teresa for welcoming me in the lab and sharing a great time inside and outside the lab, which has resulted in a long lasting friendship. Next, I would like to thank Martin, with whom I have shared not only a messy desk, but also many discussions from which I have learned a lot. Every single member of the Nevado and neighboring groups has made my time as PhD student unforgettable: I would like to thank all the current lab members, as well as Emma, Linglin, Fiona, Riccardo, Maria, Manuela and Melanie. I would also like to thank the members of the Caflisch group for all the interesting collaborations, especially Tim, with whom I have had many interesting discussions.

Regarding people not related to my research environment, I would like to start thanking my family in Zurich, my flatmates Leila and Pavlin, for sharing many unforgettable moments. I am also extremely grateful to Mikel, Amaia and Luis for all the fun, still in the distance, through skype.

Finally, I would like to thank my ‘‘ama’’, Leyre and my grandmum ‘‘Txitxus’’ for their unconditional love and support. Without them, this work would not have been possible. Last, but not least, I would like to thank Arda for being the most understanding and lovely ‘‘fistik’’ anyone could wish to have. I would have not made it this far without him.

# ZUSAMMENFASSUNG

Die Entdeckung und Entwicklung von neuen Arzneimitteln ist ein interdisziplinärer Prozess an dem viele unterschiedliche Forschungsbereiche wie Biologie, Chemie, Medizin, Kristallographie, Computer Science und noch viele weitere beteiligt sind. Die Medizinalchemie leistet durch die Entwicklung von neuen Arzneimittel und Diagnostika einen wichtigen Beitrag in der heutigen Gesellschaft.

Der erste Teil der Dissertation behandelt die Entwicklung und Synthese von Inhibitoren für EphB4 Rezeptor Tyrosin Kinase. EphB4 wird mit der Angiogenese, welche die Bildung neuer Blutgefässe beinhaltet und die Tumorzellen mit den nötigen Nährstoffen versorgen, in Verbindung gebracht. Protein Kinasen übernehmen durch die Phosphorylierung bestimmter Proteine eine wichtige Rolle bei der Zellsignalisierung. Daher ist die Hemmung der enzymatischen Aktivität durch kleine organische Moleküle ein wichtiges und weitgehend erforschtes Thema im Bereich von Drug Design. In dieser Arbeit wurden die biologischen Eigenschaften von einem EphB4 Inhibitor, welcher mittels Computer Simulation entdeckt wurde, durch die Synthese von mehreren Analoga verbessert. Deren Bindungsaffinitäten wurden mittels biochemischen und zellbasierten Tests bestimmt und in Folge dessen wurde ein Analog in einem in vivo cancer model validiert.

Im zweiten Teil der Arbeit wird die Entwicklung von neuen Liganden für Bromodomänen ausgehend von einem potentiellen micromolaren Liganden, welcher durch in silico Methoden entdeckt wurde, beschrieben. Bromodomänen sind epigenetische Leser Proteine, die eine wichtige Rolle im Bereich der Arzneimittelforschung darstellen und somit als wertvolle Zielmoleküle für die Entwicklung von neuen Arzneimittel in Betracht gezogen werden können. Durch strukturbasiertes Design von Analoga konnten nanomolare potentielle CREBBP Liganden mit einer sehr hohen Selektivität innerhalb der Familie der Bromodomänen erzielt werden. Die Untersuchung von den synthetisierten Liganden gegen einige Krebszellen zeigte, dass diese möglicherweise für eine Therapie gegen Leukämie anwendbar sind.

Im Rahmen der Dissertation wurde auch an Aktin, ein attraktives aber bis dato wenig erforschtes Zielmolekül in der Medizinalchemie, gearbeitet. Aktin ist ein Zytoskelettprotein, welches in vielen wichtigen zellulären Funktionen involviert ist und auch bei pathogenen zellulären Prozessen wie zum Beispiel Angiogenese, Zelladhäsion, Zytokinese oder Metastasen eine wichtige Rolle spielt. Mit Hilfe von neuen Berechnungsansätzen wurde eine Auswahl von vielversprechenden Molekülen, die an die ATP Bindungsstelle von Aktin binden, synthetisiert und getestet. Diese kleinen organischen Moleküle enthalten nützliche Tools um die Dynamik von Aktin zu studieren, da es ihnen möglich ist das Zytoskelett von Aktin in Zellen zu

modifizieren und die Aktin Polymerisation in vitro zu hemmen. Daher können diese Moleküle als Ansatzpunkt für die Entwicklung von weiteren potentiellen Aktin Inhibitoren dienen.

# SUMMARY

Drug discovery and development is a very challenging interdisciplinary endeavor that needs the contribution of medical doctors, biologists, chemists, X-ray crystallographers, and computer scientists, among many others, in order to be successful.

The first part of this Ph. D. thesis focuses on the development of EphB4 receptor tyrosine kinase inhibitors. EphB4 has been linked to angiogenesis, which involves the formation of new blood vessels supplying tumor cells with the necessary nutrients. Protein kinases play a key role in cell signaling by phosphorylating specific proteins and thus, the inhibition of their enzymatic activity by small organic molecules has been widely explored in drug design. In this work, the biological properties of an EphB4 inhibitor identified by computer simulations were improved by the synthesis of several analogues. Their binding affinities were characterized by an array of biochemical and cell based assays, concluding with the validation of one of the most promising derivatives in an *in vivo* cancer xenograft model.

The second part of the thesis deals with the development of novel bromodomain ligands starting from a micromolar potent *in silico* discovered hit. Bromodomain proteins are epigenetic readers that constitute an emerging topic in the field of drug discovery and are thus considered as very attractive targets for the development of novel therapeutic drugs. A careful, structure-based design of analogues resulted in the discovery of nanomolar potent CREBBP ligands with an unprecedented selectivity profile among the bromodomain protein family. Moreover, the screening of the synthesized analogues against several cancer cell lines revealed leukemia as a possible therapeutical application for the developed compounds.

The third aspect of this work deals with actin: a very attractive, but yet unexplored target in medicinal chemistry. Actin is a cytoskeletal protein that participates in many important cellular functions and has been linked to key pathogenic cellular processes such as angiogenesis, cell adhesion, cytokinesis and metastasis. A new computational approach to discover novel actin leads targeting the ATP binding site of actin resulted in the selection of promising compounds, which were synthesized and tested. The developed small organic molecules constitute valuable tools for the study of actin dynamics as they are able to modify the actin cytoskeleton in cells and moderately inhibit actin polymerization *in vitro*; thus becoming promising starting hits for the development of more potent actin binders.

# 1

# INTRODUCTION

## 1.1 Motivation

The reduction in chemistry research in pharmaceutical industry due to economic reasons has generated an important niche, creating wide opportunities for academic research groups to contribute at early stages of drug development campaigns.[1] In fact, important discoveries that shaped the pharmaceutical industry in the last century such as the isolation of penicillin and its use as antibiotic by Alexander Fleming in 1928 or the use of nucleoside analogues such as AZT as antiviral drugs, are examples of ground-breaking results stemming from academic research groups. In fact, between 1970 and 2009, there were 93 new small-molecule drug applications filled in the USA, 28 of which originated in academia.[2]

### 1.1.1 Drug discovery process

The development of drugs is a long and expensive process (Figure 1),[3] which, although not necessarily, usually starts with the identification of a target receptor that needs to be affected in order to get the desired therapeutic effect. The next step is to find an active compound that can achieve such an effect, which is referred to as a hit. Hit identification can either be done by screening a compound library composed of several thousand compounds experimentally *in vitro* (High Throughput Screening, HTS) or *in silico*. For this purpose, different types of libraries can be used, such as in house libraries, commercial libraries and targeted libraries, which are libraries of compounds compiled obeying certain specific criteria. If the studied target has known inhibitors or ligands, the library can be assembled with molecules that resemble in structure or properties to the known substrates (referred to as ''ligand-based design''). Moreover, if a crystal structure of the target is known, ''structure-based design'' can be performed, where compounds that are

complementary to the binding site are pursued.[4] In ''structure-based design'', virtual screening and *de novo* design approaches are valuable tools able to analyze the receptor shape and interaction sites.

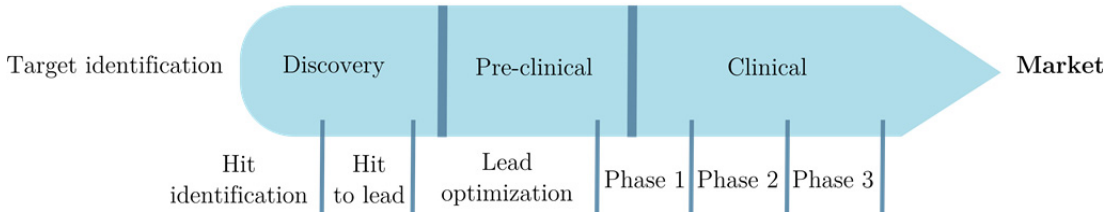



**Figure 1.** Drug discovery process.

Once the hits have been secured, *in vitro* validation, which is based on assays using isolated proteins and can easily be automated, enables the screening of thousands of compounds in a short period of time (HTS). Compounds showing a positive effect *in vitro* are further tested on cell based assays.

The hit compounds usually present modest activity, which commonly lies in the micromolar range. During hit to lead optimization dramatic changes are performed into the original chemical template aiming to identify the essential core of the lead compound and improve the potency towards the nanomolar range.[5] Structure-Activity Relationships (SAR), physico-chemical properties, chemical tractability and synthetic accessibility are thoroughly studied through the synthesis of numerous derivatives. Once the molecule's core has been established, the lead needs to be optimized by making changes in the molecule's periphery in order to obtain target affinity, selectivity, water solubility, membrane permeability and drug-like properties. At this stage, very small changes in the molecule can lead to significant effects when testing the compounds in cells or *in vivo*.

Finally, compounds showing the desired therapeutic effect on animal models (*in vivo*) are selected for clinical trials, which are divided in different phases. In the first phase safety is evaluated; during phase II the efficacy is tested; and finally, the effectiveness, benefits and possible adverse reactions are studied during the phase III.

## 1.2 Challenges

The major challenges in the development of drug candidates in an academic environment are summarized below.

- **Target selection:** The modulation of the target of choice should be translated into a desired therapeutical effect, and thus, compelling biological

evidence should be available in the literature that validates the target as therapeutically relevant. Moreover, novelty and druggability of the target are important factors to take into account.

- **Virtual screening:** Due to the inaccessibility of high throughput screening facilities in academic institutions, collaborations with bioinformatics groups are often necessary in order to carry out virtual screening campaigns to identify hits in those projects where the structure of the receptor is known.
- **Lead optimization:** Lead optimization is the major expense during pre-clinical drug development in pharmaceutical industry due to the large number of analogues that are synthesized to improve the potency of the identified hits. Organic synthesis is highly labor intensive and unpredictable, as reactions might not work as expected, yields might be poor and purifications tedious.[1] In industry, a minimum of hundreds of analogues are synthesized and tested during a successful optimization campaign. However, due to the nature of the Ph. D. work with limited time and human work force, a careful structure based design is necessary prior to synthesis, so that a reduced number of analogues are synthesized in a rational fashion.
- **Biological evaluation:** In order to asses the potency and efficacy of the synthesized analogues, several *in vitro* and cell based assays need to be established, which can potentially include protein expression and purification, X-ray crystallography, isothermal titration calorimetry (ITC), surface plasmon resonance (SPR) and differential scanning fluorimetry (DSF) among many other techniques. As a consequence, multidisciplinary expertise is required and academic collaborations with biological groups are often sought.

## 1.3 Outline of the thesis

The goal of this thesis is the design, synthesis and biophysical characterization of lead compounds against a variety of biological targets (including kinases, bromodomains and actin) that are relevant for pathologies such as cancer, inflammation and neurodegenerative diseases. Such a multidisciplinar work has been achieved thanks to collaborations with the group of Prof. Dr. Amedeo Caflisch (Computational Structural Biology, Biochemistry Institute, UZH).

Chapter 2 describes the design of kinase inhibitors based on the X-ray crystal structures of the catalytic domain of the EphA3 tyrosine kinase in complex with two inhibitor hits previously discovered *in silico*. Several derivatives were synthesized and

characterized with an array of different techniques including enzymatic assays, DSF, SPR, cell based assays and *in vivo* efficacy studies.

In chapter 3, a medicinal chemistry campaign starting from an *in silico* discovered hit targeting the CREBBP bromodomain is presented. The quest for bromodomain inhibitors as potential therapeutic tools has recently begun thanks to the growing understanding of epigenetic processes. During this work, the potency of the CREBBP hit was improved and the selectivity profile of the synthesized analogues was determined in a broad panel of bromodomain proteins. In addition, the cell growth inhibition of the analogues was characterized in different human cancer cell lines.

Chapter 4 deals with the *in silico* design, synthesis and biological evaluation of actin inhibitor hits. Actin constitutes a novel and attractive target, which has been linked to several key cellular processes, including cell division and metastasis. However, no small organic inhibitors (except complex natural products) are known to interact with actin, which makes the development of simple actin inhibitors highly valuable.

Finally, chapter 5 summarizes the work presented in this thesis and suggests the next possible steps for future research.

## 1.4 List of publications and patents

The following publications have stemmed from this work:

[I] A. Unzue, J. Dong, K. Lafleur, H. Zhao, E. Frugier, A. Caflisch, and C. Nevado, ‘‘Pyrrolo[3,2-b]quinoxaline Derivatives as Types I1/2 and II Eph Tyrosine Kinase Inhibitors: Structure-Based Design, Synthesis, and in Vivo Validation,’’ *Journal of Medicinal Chemistry*, Vol. 57, No. 15, pp. 6834–6844, 2014.

[II] A. Unzue, K. Lafleur, H. Zhao, T. Zhou, J. Dong, P. Kolb, J. Liebl, S. Zahler, A. Caflisch, and C. Nevado, ‘‘Three stories on Eph kinase inhibitors: from *in silico* discovery to *in vivo* validation,’’ *Submitted to Current Drug Targets in November 2014.*

[III] A. Unzue, J. Dong, K. Lafleur, H. Zhao, E. Frugier, A. Caflisch, and C. Nevado, ‘‘2-amino-1-phenyl-pyrrolo[3,2-b]quinoxaline-3-carboxamide derivatives,’’ *Patent application number 14162145.8. Patent number 1462. Patent pending.*

[IV] A. Unzue, M. Xu, J. Dong, L. Wiedmer, D. Spiliotopoulos, A. Caflisch, and C. Nevado, ‘‘Fragment-based Design of Selective Nanomolar Ligands of the CREBBP Bromodomain,’’ *Submitted to Angewandte Chemie in December 2014.*

[V] M. Xu, A. Unzue, J. Dong, D. Spiliotopoulos, C. Nevado, and A. Caflisch, ‘‘CREBBP bromodomain ligands discovered by docking, optimized by molecular dynamics-guided synthesis of derivatives, and validated by X-ray crystallography,’’ *Manuscript in preparation.*

[VI] A. Unzue, M. Xu, J. Dong, D. Spiliotopoulos, A. Caflisch, and C. Nevado, ‘‘Compounds, in particular for use in the treatment of a disease or condition for which a bromodomain inhibitor is indicated,’’ *Patent pending.*

## 1.5 References

1. Jorgensen, W. L. Challenges for Academic Drug Discovery. *Angew. Chem. Int. Ed.* **2012**, *51*, 11680-11684.
2. Stevens, A. J.; Jensen, J. J.; Wyller, K.; Kilgore, P. C.; Chatterjee, S.; Rohrbaugh, M. L. The Role of Public-Sector Research in the Discovery of Drugs and Vaccines. *New Engl. J. Med.* **2011**, *364*, 535-541.
3. Hubbard, R. E. *Mol. BioSyst.* **2005**, *1*, 391-406.
4. Jorgensen, W. L. Efficient Drug Lead Discovery and Optimization. *Accounts Chem. Res.* **2009**, *42*, 724-733.
5. Hann, M. M. Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Commun.* **2011**, *2*, 349-355.

# 2

# EPH TYROSINE KINASE INHIBITORS

## 2.1 Introduction

### 2.1.1 Function, classification and structure of protein kinases

Protein kinases are enzymes that catalyze the transfer of the γ phosphate in ATP to the hydroxyl group of serine, threonine or tyrosine in other proteins (Figure 1).[1]

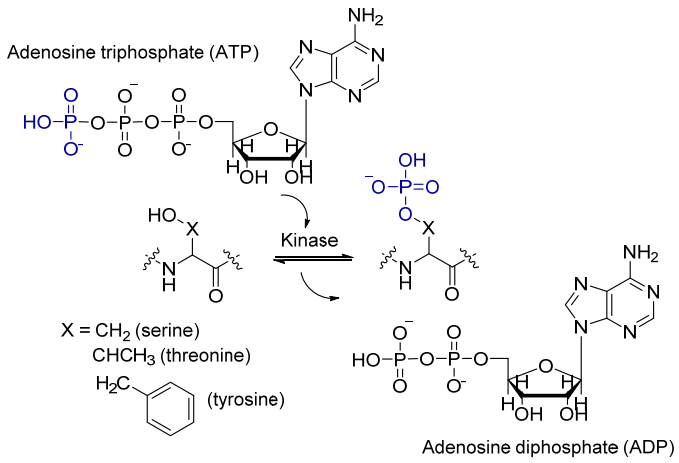



**Figure 1.** The protein kinase catalyzes the transfer of the γ-phosphate of adenosine triphosphate (ATP) to a hydroxyl-containing amino acid side chain.

Protein phosphorylation is an important post-translational modification (PTM). During protein biosynthesis, the synthesized polypeptide chains undergo PTM, which are covalent modifications that alter both their structure and function. Protein phosphorylation causes a conformational rearrangement in the protein, modifying its function and transmitting a cellular response. This process is reversible and the dephosphorylation is performed by phosphatases.[2]

Kinases are one of the largest protein families, with 518 protein kinases encoded in the human genome.[3] The so called conventional protein kinases (the most abundant ones) are divided into eight families: AGC, CAMK, CK1, CMGC, STE, TK, TKL and others (Figure 2).[3]

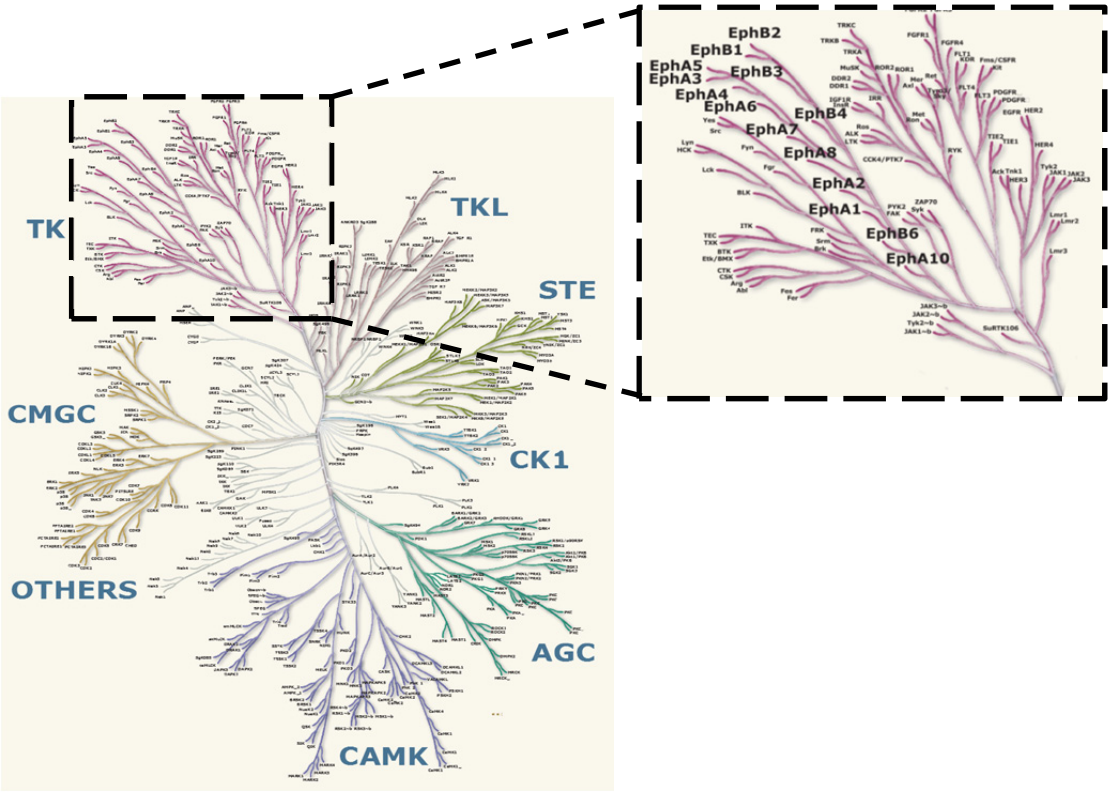



**Figure 2.** Phylogenetic tree of the conventional human protein kinases, which are divided in eight families and are further classified into subfamilies. A close-up of the tyrosine kinase (TK) family (composed of 80 kinases), relevant in this work, is shown on the right. Special attention will be given to the ephrin receptors later in this chapter (highlighted in bold).

Receptor tyrosine kinases (RTKs, Figure 2) are transmembrane proteins: the C-terminus holds the protein tyrosine kinase domain and is located in the cytoplasm of the cell, whereas the N-terminus contains the ligand binding domain and is present in the cell surface (Figure 3, A).[4] Both regions are connected by a transmembrane domain (Figure 3, A, in green). Signal molecules such as protein growth factors bind to the extracellular domain (N-terminal domain), causing two receptors to dimerize (Figure 3, B). The dimerization activates the tyrosine kinase activity and autophosphorylation of tyrosine residues in the receptor tyrosine kinase takes place (Figure 3, C). In the next step, a variety of signaling proteins bind to the autophosphorylated tyrosines. As a result, signaling proteins get phosphorylated and a cellular response is transmitted (Figure 3, D).

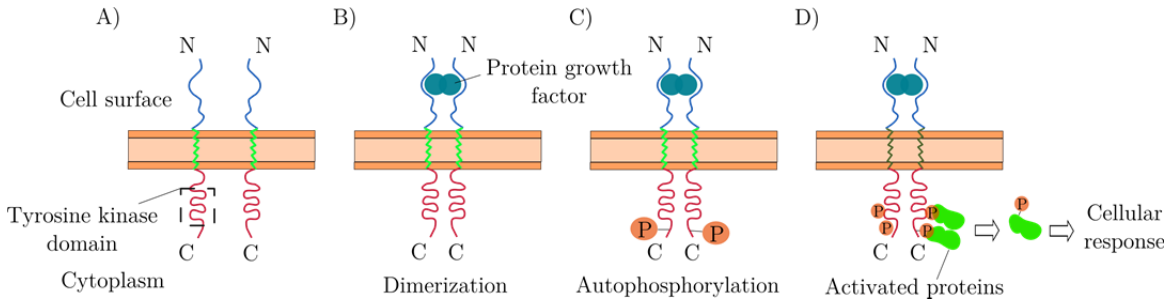



**Figure 3.** Dimerization and activation of receptor tyrosine kinases.[4]

Tyrosine kinase domains (Figure 3, A) are composed of about 250 aminoacids and can be divided into a carboxy terminal (C-terminal, Figure 4, in blue) and an amino terminal (N-terminal, Figure 4, in magenta) lobes or subdomains. The N-terminal subdomain consists of a barrel of β sheets and one conserved α helix, whereas the C-terminal lobe is primarily composed of α helices and accommodates the peptide substrate (Figure 4, shown as sticks in green).

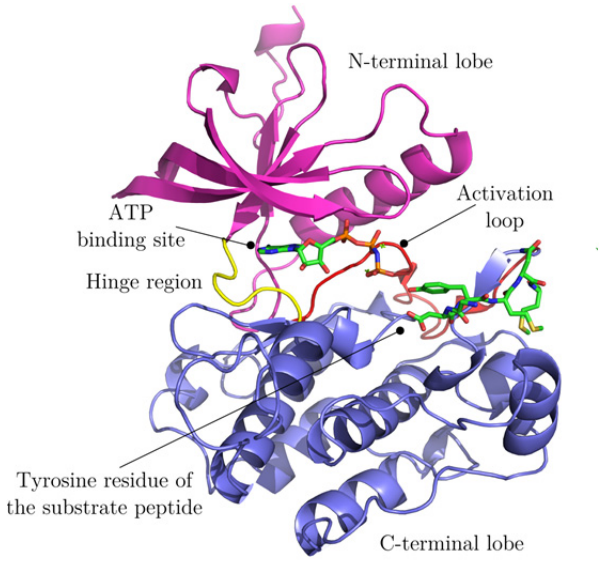



**Figure 4.** Structure of the insulin receptor tyrosine kinase. The N-terminal lobe is shown in magenta, the hinge region in yellow and the C-terminal lobe in blue. The receptor kinase domain is shown with an ATP analogue and its peptide substrate bound to the C-terminal lobe (PDB code 1IR3).

Both subdomains are connected by the so called hinge region (Figure 4, in yellow), where the ATP and a $Mg^{2+}$ ion bind.[5, 6] All kinases possess a flexible activation loop (Figure 4, in red) that can adopt different conformations and regulates the kinase activity. The DFG is a sequence of three amino acids, Asp-Phe-Gly, located within the activation loop. This segment can adopt different conformations depending on the orientation of the phenylalanine and aspartate residues. In the so called DFG-in conformation the enzyme is

catalytically active (usually phosphorylated), whereas in the DFG-out conformation the activation loop prevents the substrate from binding and the enzyme is dysfunctional.[7]

The ATP binding pocket is divided into different subpockets as indicated in Figure 5.[8] The gatekeeper residue (threonine in Figure 5, but variable among kinases) determines the size of the hydrophobic pocket, and is therefore of high importance for the design of selective kinase inhibitors. The adenosine moiety of ATP forms two hydrogen bonds within the hinge region, a binding mode which is conserved in all kinases (with the exception of PIM kinases). As the backbone of the hinge region is identical in almost all kinases, the design of selective kinase inhibitors has proven to be very challenging. In the next section, an overview of the different kinase inhibitor ligands developed to overcome the above mentioned selectivity issue is given.

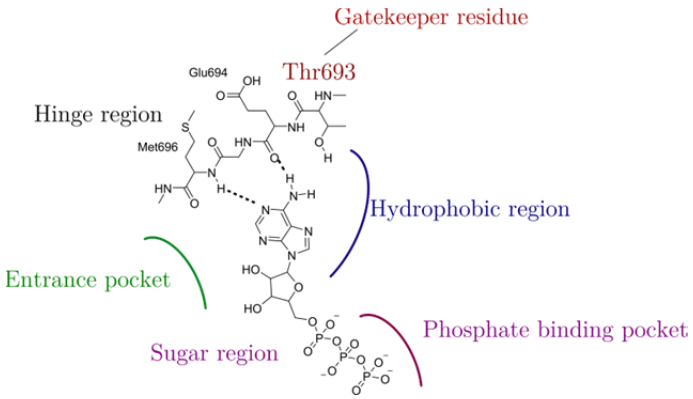



**Figure 5.** Schematic representation of the binding site of an ATP bound kinase.

### 2.1.2 Protein kinases as drug targets and their binding mode

Protein kinases play a key role in cell biology, metabolism, cell migration, apoptosis, proliferation and differentiation. The activity of protein kinases has been linked to a wide variety of therapeutic areas such as oncology,[7] neurology,[9] immunology,[10] cardiology[11] and infectious diseases.[12] Among the mentioned therapeutic areas cancer has been notably addressed. As certain protein kinases are over-expressed in many cancer cells, kinase inhibition has been seeked for the development of more selective anti-cancer agents.

Since the first ATP competitive kinase inhibitor staurosporine in 1986,[13] several small molecule kinase inhibitors have been approved by the U.S. Food and Drug Administration (FDA) for cancer treatment (Figure 6).[14] Kinase inhibitors are classified into covalent (Figure 5, in red) or non-covalent inhibitors, and the latter are further divided into different types depending on their binding mode: type I-III.

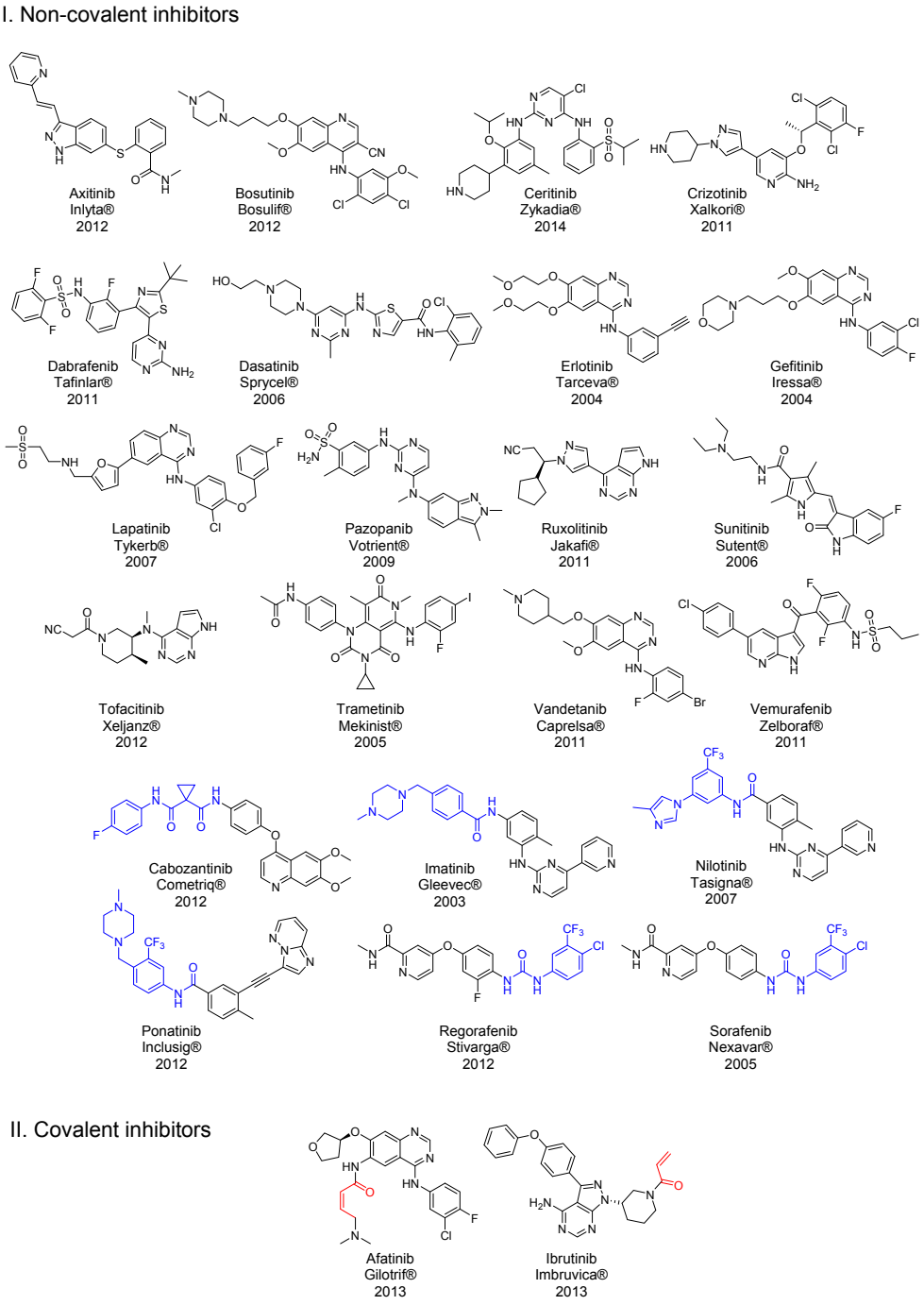



**Figure 6.** Kinase inhibitors approved by the FDA.[14] Structural motifs that belong to type II inhibitors and bind to the DFG-out pocket are highlighted in blue, and the electrophile moieties that react in a covalent fashion with the protein are shown in red.

### I. Non-covalent kinase inhibitors

The first ever developed kinase inhibitors have been classified as type I, as they bind within the ATP binding site mimicking the purine ring of ATP and are thus, direct ATP competitors. They recognize the activated state of the kinase, which is characterized by an open conformation of the activation loop (DFG-in, Figure 4, in red).[15]

Type I inhibitors usually consist of a heterocyclic ring system that forms one to three hydrogen bonds with the hinge region mimicking the adenine of ATP. In addition, they can form extra interactions in adjacent hydrophobic regions. An schematic representation of a type I inhibitor pharmacophore is shown in Figure 7, A.[7, 15-17]

The main drawback of type I inhibitors is the low kinase selectivity, due to the high degree of sequence and structural similarity in the ATP binding site in all kinases. This translates in undesired off-target side effects and toxicity.[18, 19] A very well-known FDA approved drug for the treatment of chronic myelogenous leukemia, dasatinib, is a clear example of a type I kinase inhibitor (Figure 7, B-D).[20] Dasatinib binds to the active form (DFG-in) of the kinase and forms two hydrogen bonds with the hinge region and a third one with the gatekeeper residue. Due to the low sequence variation of the hinge region among different kinases, dasatinib presents a broad selectivity profile (Figure 6, D).

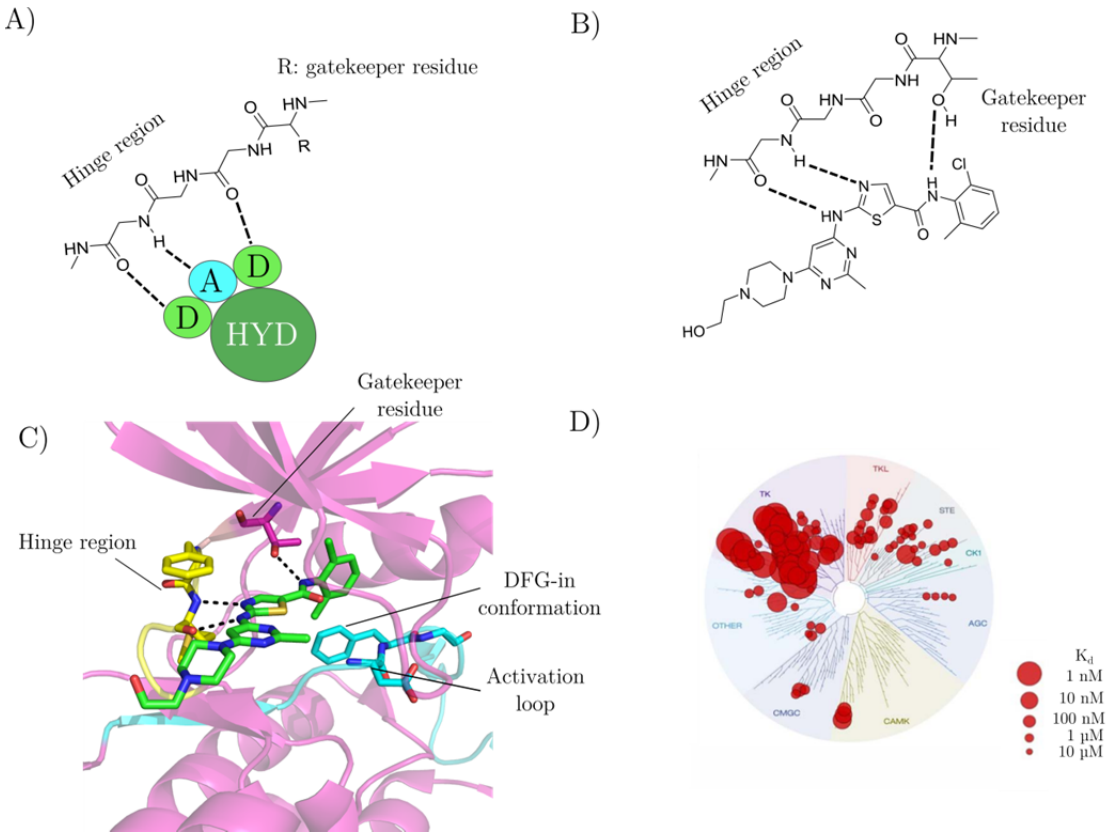



**Figure 7.** A) Type I inhibitor pharmacophore. The hydrogen bonds with the hinge region are shown as dashed lines. Hydrogen bond donors are indicated as D, hydrogen bond acceptors as A and hydrophobic moieties are labeled HYD.[15] B) Kinase bound to dasatinib in its active form as a 2D representation. C) 3D representation of the Abl1 kinase in complex with dasatinib (PDB code 2GQG). The hinge region is shown in yellow, the activation loop in light blue and the ligand, dasatinib, in green. The DFG residues are shown as sticks in blue. D) Selectivity profile of dasatinib against 442 kinases. The dissociation constant, $K_d$, is shown as red spheres.[18]

In contrast to type I inhibitors, type II ligands do not compete with ATP. They interact with an inactive conformation of the kinase (unphosphorylated), which contains a different conformation of the activation loop (Figure 4). In a typical type II inhibition mode, the activation loop is in the DFG-out conformation. A new pocket arises from the change in conformation of the phenylalanine residue (located 10 Å away)[16] in the DFG loop (activation loop), and as a result, neither ATP nor protein substrates can bind and the protein is inactive (Figure 8).[15]

While still interacting in the hinge region forming one to three hydrogen bonds, type II inhibitors exploit new interactions within the newly formed lipophilic pocket in the back of the protein (Figure 8, B).[15] The extra interactions formed in the DFG-out pocket and the lower sequence conservation on the inactive kinase conformation among different kinases provide type II inhibitors a higher degree of selectivity.[19, 21, 22] Moreover, not all kinases are able to adopt such a conformation.[16] In fact, kinases that can adopt a DFG-out conformation often have a small gatekeeper residue, usually threonine.[23] Another advantage offered by type II inhibitors is the slower dissociation constants ($k_{off}$) and longer residence times, which makes them valuable tools for drug development.[24, 25] The main drawbacks of type II kinase inhibitors are their high molecular weight (which influences cell permeability and ligand efficiency)[26, 27] and a higher vulnerability to mutations. A mutation in a residue not involved in the catalytic activity of the kinase or ATP binding could still impair the binding of a type II compound, decreasing its efficacy. On the other hand, a mutation that prevents a type I inhibitor from binding would, at the same time, block the binding of ATP (since they share the same binding site) and result in a dysfunctional kinase.[15, 28]

A clear example of a type II inhibitor is imatinib (Figure 8, B),[29] which was the first kinase small molecule inhibitor to reach the market and is used for the treatment of chronic myelogenous leukemia, acute lymphoblastic leukemia, aggressive systemic mastoytosis and gastrointestinal stromal tumor.[14] The pyridine nitrogen of imatinib forms one hydrogen bond with the hinge region, and interacts with the glutamic acid of the α-helix and the aspartic acid of the DFG-out pocket via hydrogen bonds with its amide linker, a conserved interaction throughout type II inhibitors (Figure 8, A, B, C).[15] Imatinib presents a more specific selectivity profile (Figure 8, D) compared to dasatinib (Figure 7, D), as generally seen for type II inhibitors. The main strategy for the development of imatinib and other type II inhibitors is known as hybrid-design,[30] where a type I binder is extended with a hydrophobic tail that targets the DFG-out pocket and forms van der Waals interactions in that region. The mentioned hydrophobic tail is usually linked to the type I motif by amides or ureas, which provide a hydrogen bond donor that interacts with the glutamic acid and a hydrogen bond acceptor that interacts with the aspartic acid residue of the activation loop (Figure 6, shown in blue).[30]

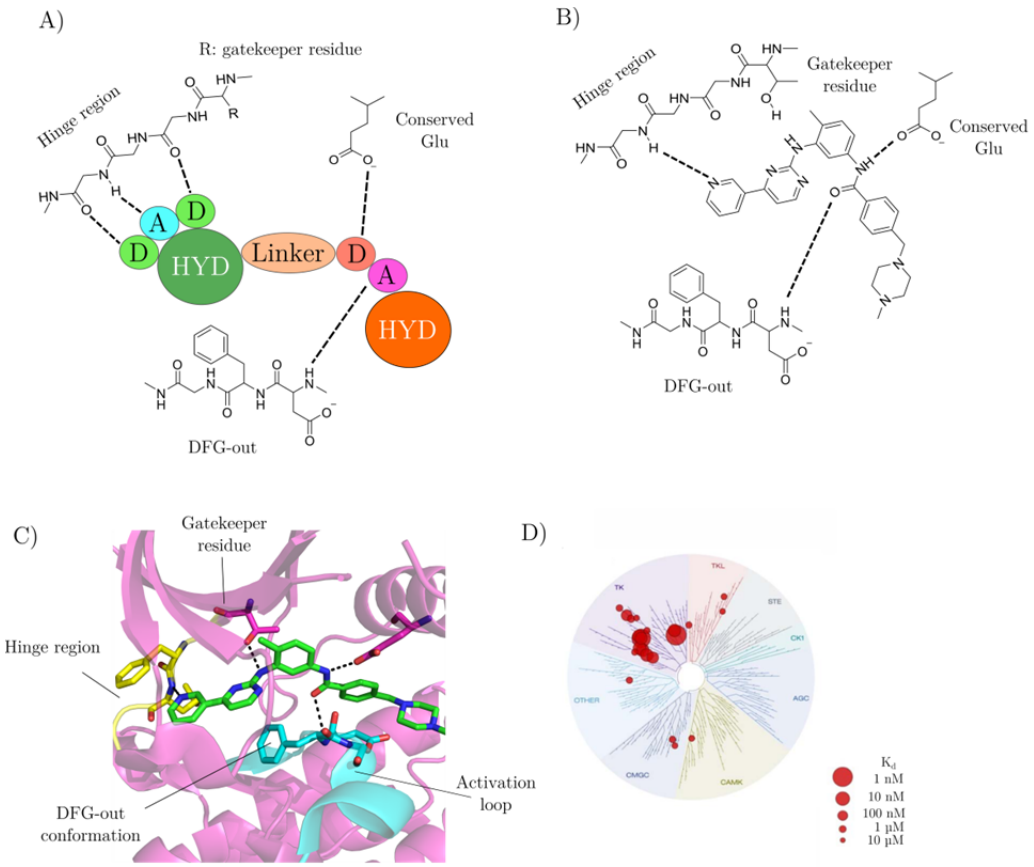



**Figure 8.** A) Type II inhibitor pharmacophore. The hydrogen bonds with the hinge region, the glutamic acid residue located in the α-helix and the activation loop bearing a DFG-out conformation are shown as dashed lines. Hydrogen bond donors are indicated as D, hydrogen bond acceptors as A and hydrophobic moieties are labeled HYD.[15] B) Kinase bound to imatinib in its inactive DFG-out conformation as 2D representation. C) 3D representation of the Abl1 kinase in complex with imatinib (PDB code 2HYY). The hinge region is shown in yellow, the activation loop in light blue and the ligand, imatinib, in green. The DFG residues are shown as sticks in blue. D) Selectivity profile of imatinib against 442 kinases. The $K_d$ is shown as red spheres.[18]

A third class of kinase inhibitors, type $I_{1/2}$ has been defined, whose binding mode lies in between type I and type II. They bind to the kinase in the active DFG-in conformation, but they extend their interactions to the hydrophobic back pocket (Figure 5) establishing the characteristic hydrogen bonds present in type II inhibitors (Figure 9, A-C). The size and nature of the backpocket are provided by the gatekeeper residue, which contributes towards a selectivity gain. Potent and selective EphB4 kinase $I_{1/2}$ inhibitors have previously been developed in our group in collaboration with Prof. Amedeo Caflisch (Biochemistry Institute, UZH).[8, 31-33] As an example, a nanomolar potent xanthine type $I_{1/2}$ inhibitor (**I**) is shown in Figure 9, B.

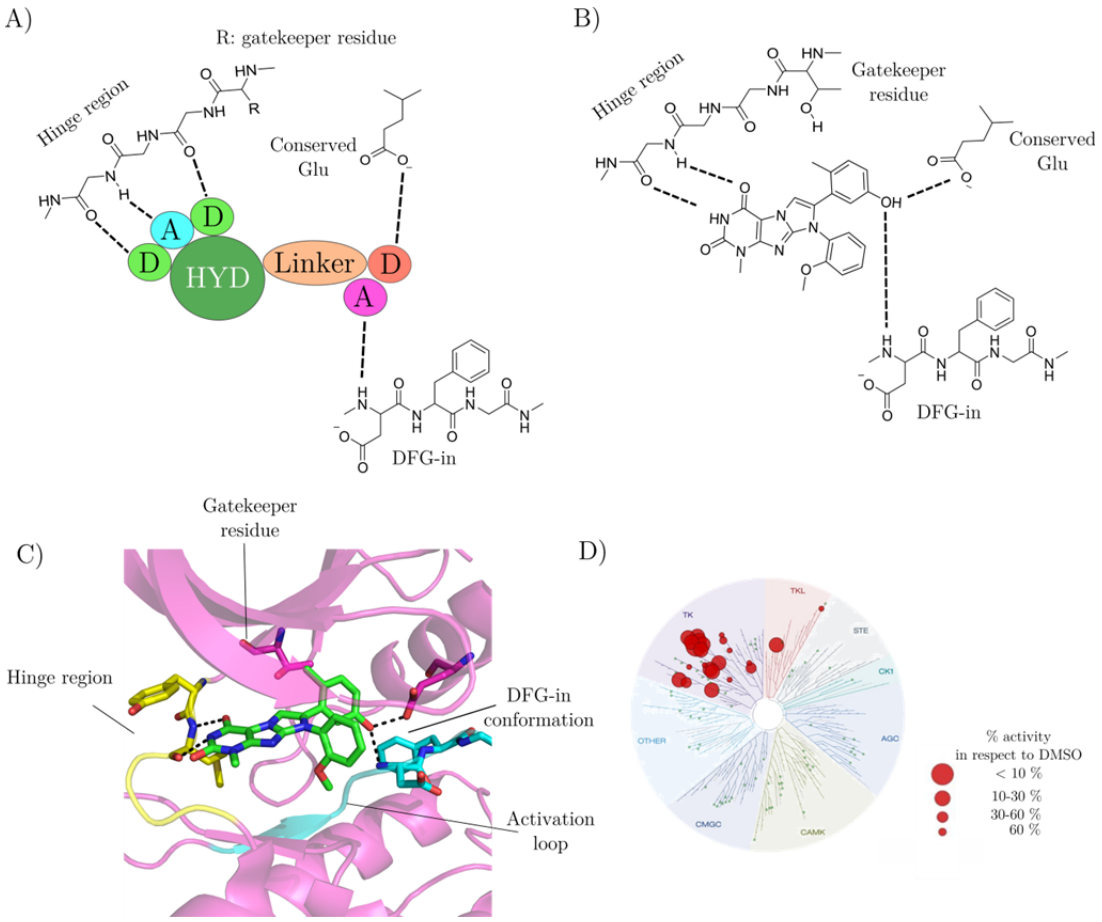



**Figure 9.** A) Type $I_{1/2}$ inhibitor pharmacophore. The hydrogen bonds with the hinge region, the glutamic acid residue located in the α-helix and the activation loop bearing a DFG-in active conformation are shown as dashed lines. Hydrogen bond donors are indicated as D, hydrogen bond acceptors as A and hydrophobic moieties are labeled HYD.[15] B) Kinase bound to the in house developed EphB4 inhibitor **I** in its active DFG-in conformation as a 2D representation. C) 3D representation of the EphA3 kinase in complex with compound **I** (PDB code 4GK2). The hinge region is shown in yellow, the activation loop in light blue and the ligand, **I**, in green. The DFG residues are shown as sticks in blue. D) Selectivity profile of **I** against 143 kinases.[31]

The pyrimidine ring of ligand **I** is involved in two hydrogen bonds with the hinge region, whereas the phenol is located in the hydrophobic back pocket. The hydroxyl group of the aromatic ring donates a hydrogen bond to the conserved aspartic acid and accepts a hydrogen bond from the backbone of the DFG loop in the DFG-in conformation (Figure 9, B, C).[31] The selectivity profile of compound **I** shows strong inhibition of a relatively small fraction of the human kinome, most of them having a threonine residue as gatekeeper.[31, 32]

Type III inhibitors are allosteric inhibitors that bind outside of the catalytic domain (ATP-binding site) of the kinase, but in regions where the activity of the enzyme can be modulated. Since these regions are usually unique to the kinase, higher selectivities can be obtained with type III binders. A recent example of an allosteric inhibitor is GNF-2 and its

analogue GNF-5, which were developed by Novartis and bind in an adjacent pocket to the ATP binding site of the Abl kinase (Figure 10).[34, 35] Only a few examples of type III kinase inhibitors have been reported so far,[23] but there is an emerging interest in the area due to successful results obtained with the treatment of GNF-5 in combination with the type I and II inhibitors, dasatinib and nilotinib (Figure 6), in an *in vivo* leukemia model.[34]

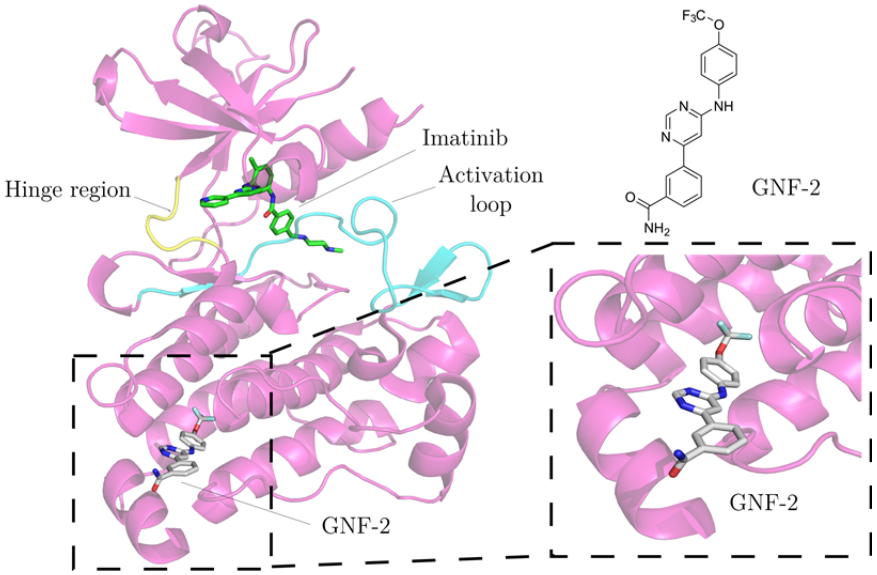



**Figure 10.** Imatinib and the allosteric inhibitor GNF-2 in complex with Abl kinase (PDB code 3K5V).

### II. Covalent kinase inhibitors

The last class of kinase inhibitors, covalent inhibitors, bear electrophiles such as good leaving groups or Michael acceptors that form covalent bonds with cysteine or other nucleophilic residues in the ATP binding pocket.[36, 37] As a result, the inhibitor irreversible blocks the ATP binding side. The main disadvantage of covalent inhibitors is their high toxicity. However, higher potency is achieved due to prolongued inhibition and off-target effects are usually minimized.[23, 30] Ibrutinib (Figure 6) is a covalent Bruton's tyrosine kinase inhibitor approved by the FDA in 2013 for the treatment of leukemia and lymphoma.[38] The double bond of the inhibitor is nucleophilically attacked by a cysteine residue located in the ATP binding site of the kinase, forming a C-S covalent bond. Despite their higher potency and minimized side effects, covalent inhibitors have only emerged recently, are less abundant and have been barely studied in comparison with non-covalent inhibitors.[16]

### 2.1.3 Eph receptor tyrosine kinases

Eph receptor tyrosine kinases are the largest receptor tyrosine kinase family (Figure 2).[39] There are nine EphA receptors, which bind to five ephrin-A ligands, and five EphB receptors that bind to three ephrin-B ligands. Eph receptors and ephrin ligands are membrane bound molecules that require cell contact in order to interact with each other.

An in detail representation of the EphB receptor is shown in Figure 11. Eph receptors are composed of a ligand-binding domain located at the end of the extracellular domain (N-terminal domain), which is directed towards the inner part of the cell by a linker: a cysteine rich region and two glycoproteins (fibronectins). The cysteine rich regions and fibronectins do not participate in ligand binding: the former stabilize the ligand-receptor complex and the latter play a structural function maintaining a correct distance between the two cells. Inside the cell, the receptor is composed of the so called juxtamembrane region, the kinase domain and two other domains (sterile-α-domain and PDZ). The ligands of the Eph receptors, ephrins, are bound to the membrane of a second cell (Figure 11).

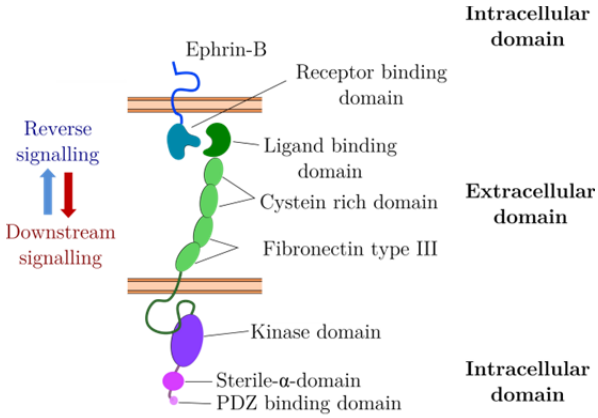



**Figure 11.** Structure of Eph receptor.[8, 40] The N-terminal domain is composed of a ligand binding domain, a cysteine rich domain and two fibronectin type III repeats. The kinase domain is located inside the cell, followed by a sterile-α-domain and a PDZ binding domain. Ephrin-B is located in another cell and consists of the receptor binding domain and a cytoplasmatic domain.

The Eph-ephrin system is characterized by its bi-directional signaling, meaning that they can downstream signal into the Eph containing cell and upstream into the ephrin containing cell.[41] The bi-directional signaling enables cell-cell communication, cell migration, cell repulsion and adhesion to the extracellular matrix among other processes.

The cell-cell interaction that occurs through Eph-ephrin interaction leads to a variety of cellular responses. The cytoskeleton is greatly affected by such interaction and changes in actin organization, cell shape, adhesion, repulsion, migration, differentiation and proliferation are promoted.[42] As a consequence, Eph-ephrin interactions modulate several

processes of the nervous system, where they are highly expressed, such as axon guidance during neural development, development of the visual system and vascular development during embryogenesis.[43]

Eph receptors and ephrins are also involved in the growth, maturation and remodeling of blood vessels. EphB4-ephrinB2 interaction in particular plays a prominent role in the cardiovascular development and it regulates vascularization of malignant tumors. In fact, EphB4 overexpression has been linked to several tumor types such as breast, prostate, bladder, uterus, melanoma and ovarian cancer among others.[11, 44-47] EphB4 has also been associated to angiogenesis, tumor growth and metastasis, making it a valuable target for therapeutical applications. However, the role of EphB4 in cancer remains a matter of debate. On the one side, EphB4 is known to promote tumor angiogenesis and growth when it is expressed on tumor cells that interact with ephrin-B2 located in the endothelium.[48] On the other side, EphB4 can either promote or suppress tumor growth depending on the tumor type and model.[45, 49-52]

Several anti-cancer strategies that address the EphB4-ephrinB2 interaction have been developed. Soluble fusion proteins of the extracellular domain of EphB4 that antagonize EphB4-ephrinB2 interaction could inhibit tumor growth and angiogenesis in tumor xenograft models, confirming the importance of EphB4 in oncology.[53, 54] Moreover, monoclonal antibodies and peptides that block EphB4-ephrinB2 interaction have also been described.[55, 56] Surprisingly, just a few nanomolar potent Eph kinase inhibitors have been disclosed (Figure 12),[57-66] which present different binding modes:[15] type I (such as **II**[57] and **III**,[58] which are ATP competitors that recognize the activated state of the kinase), type II (e.g **VI**[63] and **VIII**[64]) that interact with an inactive conformation of the kinase and generally offer better selectivity profiles) and type I $I_{1/2}$ inhibitors (e.g. **IV**[60]) that bind to the kinase in the active DFG-in conformation, but extend their interactions to the hydrophobic back pocket establishing the characteristic hydrogen bonds present in type II inhibitors.[15]

The first reported efficacy data with an EphB4 inhibitor was conducted with the selective inhibitor NVP-BHG712 (**VIII**),[64] which was shown to inhibit VEGF driven angiogenesis *in vivo*.[64] Several of the FDA approved kinase inhibitors, such as dasatinib,[67] bosutinib, vandetanib, crizotinib and nilotinib,[68-70] as well as other kinase inhibitors[71-75] are also able to inhibit EphB4 among other kinases.

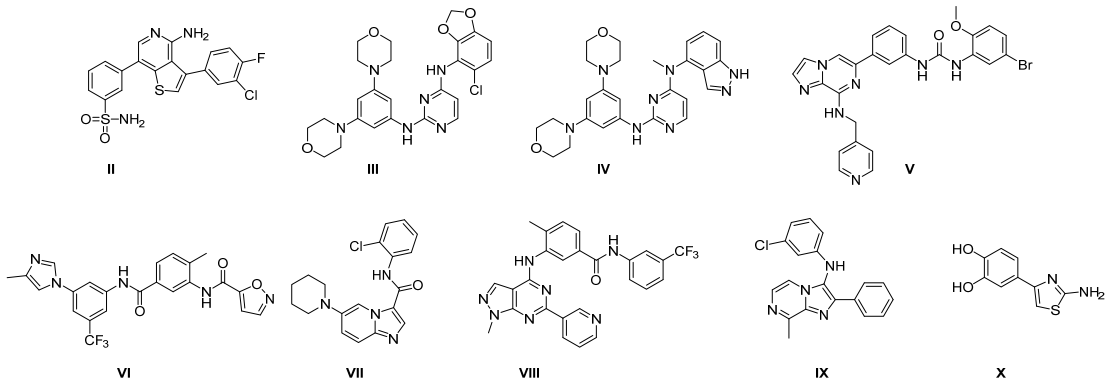



**Figure 12.** Previously reported nanomolar potent small organic molecule ephrin kinase inhibitors.[57-66]

Our groups at the University of Zurich have greatly contributed to the discovery, optimization and characterization of several selective and potent EphB4 inhibitors. On the outset of these discoveries high throughput docking techniques were applied. Specifically, we have developed an efficient computational method called ALTA (Anchor-based Library Tailoring) that generates a focused library and prioritizes molecular fragments according to their binding energy.[76] As a first step, millions of commercially available molecules are automatically decomposed into rigid fragments, which are docked in the binding site of the protein. Fragments with the most favorable binding free energy (also called anchors) are used to retrieve the molecules which they originated from. These molecules are then submitted to flexible ligand docking. Finally, compounds are ranked according to their interaction energy taking into account desolvation effects, visually inspected and chosen for testing. This *in silico* fragment based approach has resulted, after medicinal chemistry optimization, in the nanomolar potent xanthine **I**[31, 32] and pyrimidoisoquinolinone **XI** (with an $IC_{50}$ value of 9.4 μM in *in vitro* FRET assays)[33] (Figure 13).

We have also successfully applied semi-empirical quantum mechanical (QM) calculations for rapidly screening compound poses generated by high throughput docking.[77] A total of 2.7 million of commercially available compounds were docked in the ATP binding site of EphB4 generating *ca.* 100 million poses. Their binding energy was estimated by calculating the formation enthalpy in the presence of small molecules (called QM probes) emulating functional groups in the hinge region and at the entrance of the binding pocket. This study resulted in the discovery of 2-formamido-4-phenylthiophene-3-carboxamide **XII** (Figure 13), which has an $IC_{50}$ value of 5.9 μM for EphB4 in a FRET enzymatic assay. Importantly, the binding mode was confirmed by X-ray crystallography of a close derivative of **XII** in complex with EphA3 (PDB code 4TWO).[78]

Further high throughput docking studies were performed using an inactivated DFG-out conformation of the kinase obtained by explicit solvent molecular dynamic (MD)

simulations.[79] This campaign yielded a series of 5-(piperazine-1-yl)isoquinoline derivatives including inhibitor **XIII** (Figure 13) as novel type II inhibitors that present low micromolar affinities for EphA3 and unphosphorylated Abelson tyrosine kinase (Abl1).[79]

In another high throughput docking campaign, an evaluation of the hydrogen bonding penalty upon ligand binding was integrated to the binding energy calculation of docked molecules.[80] This study resulted in the discovery of three novel Eph inhibitor scaffolds: triazolo[4,3-a]quinazolines (**XIV**), phenylquinolines (**XV**) and quinoxalines (**XVI**, Figure 13).

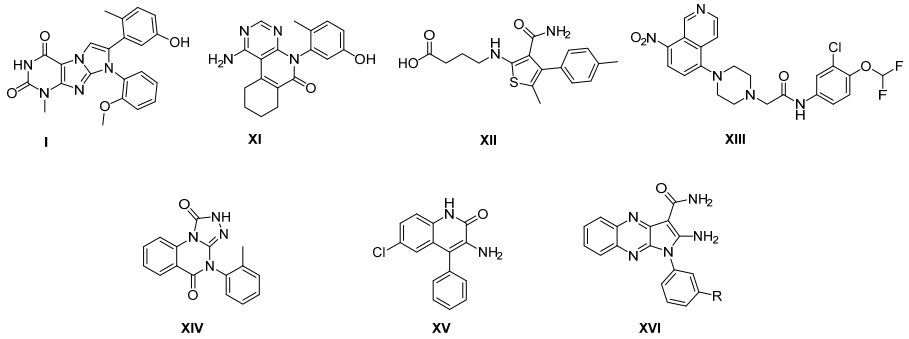



**Figure 13.** EphB4 tyrosine kinase inhibitors identified and optimized (**I**, **XI**) in the Caflisch and Nevado groups.[33, 76, 77, 79, 80]

In the next section, the development of the quinoxaline **XVI** scaffold into potent type $I_{1/2}$ and II kinase inhibitors is presented.

## 2.2 Pyrrolo[3, 2-b]Quinoxaline Derivatives as Types I1/2 and II Eph Tyrosine Kinase Inhibitors

As seen above, our groups have recently focused on the *in silico* design, synthesis and computational-aided optimization of potent and selective receptor tyrosine kinase inhibitors. Successful campaigns have yielded single-digit nanomolar EphB4 inhibitors whose potential antiproliferative activities have been characterized by cellular assays.[31-33] Furthermore, the predicted binding mode could also be confirmed by X-Ray diffraction analysis of their complexes with EphA3. Given the critical role of Eph receptors and (Eph)–ephrin signaling in tumor growth and progression,[81] a subset of these compounds are currently being pursued towards pre-clinical development.

In the next sections we describe a multidisciplinary campaign towards the design of novel and potent, type $I_{1/2}$ and II tyrosine kinase inhibitors based on the crystal structure of two type I inhibitors. The parent pyrrolo[3,2-*b*]quinoxaline scaffold **XVI** was decorated with characteristic functional groups present in previously successful type II binders, thus

speeding up the hit to lead optimization campaign. The binding kinetics of the low-nanomolar derivatives **11d** (type I1/2) and **7m** (type II) were characterized by surface plasmon resonance (SPR) measurements. Extensive profiling by biochemical (competition binding) and cellular assays, together with pharmacokinetic measurements in mice resulted in the prioritization of inhibitor **11d** for final validation *in vivo* by a human breast cancer xenograft.

### 2.2.1 Crystal structures of type I inhibitors A and B with EphA3

#### I. Docking validation by X-Ray diffraction analysis of binding complex

Recently, we reported the discovery of two type I EphB4 inhibitors **A** and **B** by automated docking.[80] The *in silico* predicted binding mode of these molecules is confirmed here by X-ray diffraction analysis of the catalytic domain of EphA3 in complex with both **A** and **B** (Figure 14). The pyrrolo[3,2-*b*]quinoxaline scaffold occupies the ATP binding site with the phenyl substituent nestled into the so-called hydrophobic pocket. The amino substituent at position 2 of the pyrrole ring is involved in a bifurcated hydrogen bond with the side chain hydroxyl of the Thr693 gatekeeper and the backbone carbonyl of Glu694. Furthermore, in the structure with inhibitor **A**, the amide substituent at position 3 of the pyrrole ring is optimally oriented for two hydrogen bonds with the backbone polar groups of Met696 so that **A** forms a total of three hydrogen bonds with the backbone of the hinge region. Only two hydrogen bonds with the same region are observed for inhibitor **B** due to the ethoxy-propyl substitution at the nitrogen of the amide whose *trans* configuration prevents it to act as donor to the carbonyl group of Met696. The lack of this hydrogen bond is consistent with the about 10-fold weaker affinity of inhibitor **B** with respect to **A** ($IC_{50}$ of 300 nM for EphB4)[80] which might also originate, at least in part, from the different substituents of the phenyl ring in the hydrophobic pocket, i.e., -$OCH_3$ and -Cl in **A** and **B**, respectively.

#### II. Design of type I1/2 and type II derivatives based on the X-ray crystal structure of the type I inhibitors A and B

Based on our previous experience[31, 32] and earlier reports towards the synthesis of potent type I kinase inhibitors,[15] several modifications within the phenyl ring were designed in order to fine tune the interactions of the quinoxaline inhibitors with the threonine gatekeeper residue (Thr693) of EphB4. Due to the rather limited space around the phenyl group revealed by the binding modes of **A** and **B**, the introduction of small substituents was envisioned, including the incorporation of a methyl and a hydroxyl group at positions

2 and 5, a combination that had been successfully exploited in our previous studies developing nanomolar inhibitors of EphB4.[32]

In addition, the binding modes of inhibitors **A** and **B** suggested the possibility to extend the quinoxaline scaffold into the allosteric binding site by substitution of the phenyl ring, resulting in type II inhibitors. Our design campaign targeting the back pocket of the kinase relied on the introduction of a variety of substituents, some of which are present in type II inhibitors. The so-called "tail" moiety of type II inhibitors, located within the allosteric binding site, is characterized by a hydrogen bond donor-acceptor pair (usually an amide or an urea), linked to a hydrophobic substituent that occupies the newly formed pocket in the DFG out conformation of the kinase.[7, 16, 30] Amide, urea and thiourea linkers were incorporated onto the quinoxaline scaffold and attached to a *m*-$CF_3$-phenyl moiety present in some type II kinase inhibitors, in analogy to AAL993,[82] sorafenib,[83, 84] and GNF-5837 [85] (**7b-d**, **7g-i**, **7l**, **7m**, **12n**). A cyclopropyl ring, a common motif in p38α isoform selective kinase inhibitors,[86-88] was also introduced (**7a**, **7f**, **7k**). In an effort to increase the hydrophobic interactions within the allosteric binding site, 4-methyl imidazole was added aiming to mimic the well-known drug nilotinib[89] (**7e**, **7j**). In order to increase the solubility, methyl or fluoro substituents were placed in *ortho*- relative position (**7c-e**, **7h-j**), therefore distorting the planarity of the molecules.[90] The results in Table 1 underline that although the identification of substituents successfully binding the allosteric pocket based on known inhibitors was not *a priori* obvious, this was nonetheless an efficient strategy to obtain potent and selective type II kinase inhibitors in a time- and cost-effective manner.

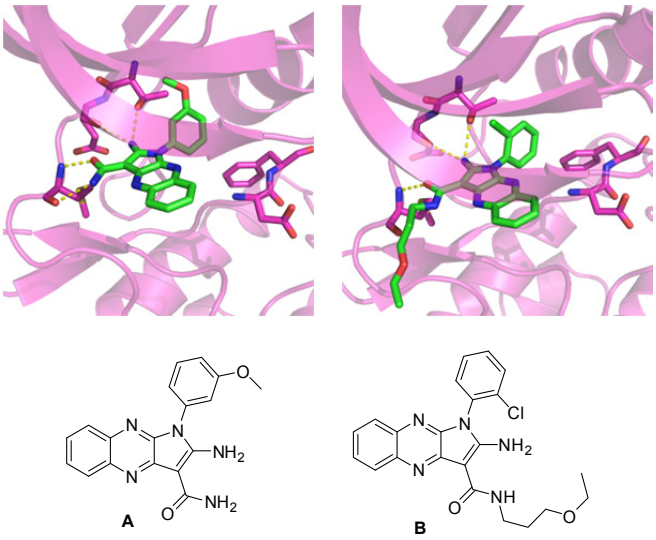


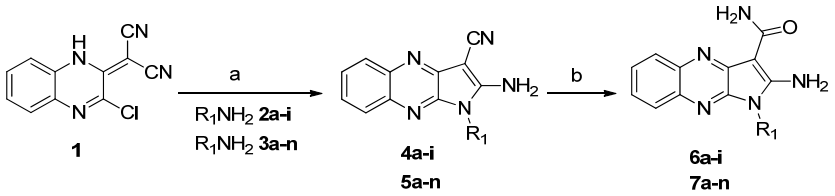

**Figure 14.** Crystal structures of the catalytic domain of the tyrosine kinase EphA3 in complex with the high-nanomolar inhibitors **A** (left, PDB code 4P4C) and **B** (right, PDB code 4P5Q). The ATP binding site of the EphA3 kinase is shown in magenta ribbons while the side chains mentioned in the text and the inhibitors are shown by sticks (with carbon atoms in magenta and green, respectively).

## 2.2.2 Characterization of new type $I_{1/2}$ and II inhibitors

### I. Synthesis

The synthesis of 1*H*-pyrrolo[2,3-*b*]quinoxaline derivatives **6** and **7** is shown in Scheme 1. Compound **1** was prepared according to previously reported procedures by condensation of commercially available 2,3-dichloroquinoxaline with malononitrile in the presence of sodium hydride.[91, 92] The substitution of the chlorine at position 3 with commercially available anilines **2a-i** followed by cyclization afforded intermediates **4a-i**.[93] The reaction with synthetically prepared anilines **3a-n** delivered tricyclic intermediates **5a-n** (Scheme 1).

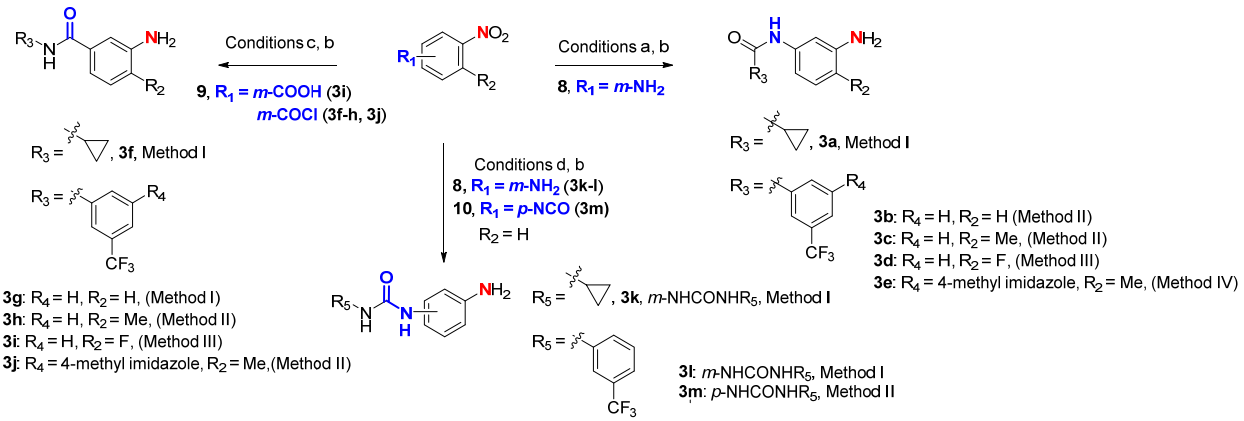
**Scheme 1**[a]




[a]Reagents and reaction conditions: a) Protocol I: $R_1NH_2$ **2a-i**, **3a-d**, **3f-i**, **3k-l**, **3n** (1-4 eq.), EtOH:Toluene, 1:1, 130-160 °C, 4-12 h, 30-79%. Protocol II: $R_1NH_2$ **3e**, **3j**, **3m** (1-1.2 eq.), DMF, 80°C, 12 h, 21-61%. b) $H_2SO_4$, 25 °C, 30 min, 7-99%.

The preparation of the non-commercially available anilines **3a-m** used in the above mentioned condensation reaction has been summarized in Scheme 2. Anilines **3a-e**, bearing a N-amide group in *meta* relative position, were obtained by condensation of 3-nitro-substituted anilines **8** with the corresponding readily available benzoyl chlorides ($R_3COCl$) or benzoic acids ($R_3COOH$)[94] using different protocols (Conditions a, Methods I-IV). Reduction of the nitro group with either $SnCl_2$ or Pd/C and $H_2$ (Conditions b) furnished the desired intermediates. Anilines **3f-j** were prepared by condensation of the 3-nitro-substituted benzoic acid or benzoyl chlorides **9** with the corresponding anilines[95] under reaction conditions c. Reduction of the nitro group with $SnCl_2$ or Pd/C and $H_2$ delivered the corresponding anilines in 30-77% overall yields. Finally, anilines **3k-m**, bearing urea moieties at *meta*- and *para*- relative position, were prepared by condensation of aniline **8** with the corresponding isocyanate (for **3k-l**) or by condensation of the isocyanate **10** with the respective commercially available aniline under reaction conditions d to give **3m**. Reduction of the nitro group under the conditions described above furnished the corresponding aniline-containing ureas **3k**, **3l** and **3m** in 42, 85 and 41% yields respectively. Aniline **3n** was prepared according to a previously reported procedure.[96]

**Scheme 2**[a]

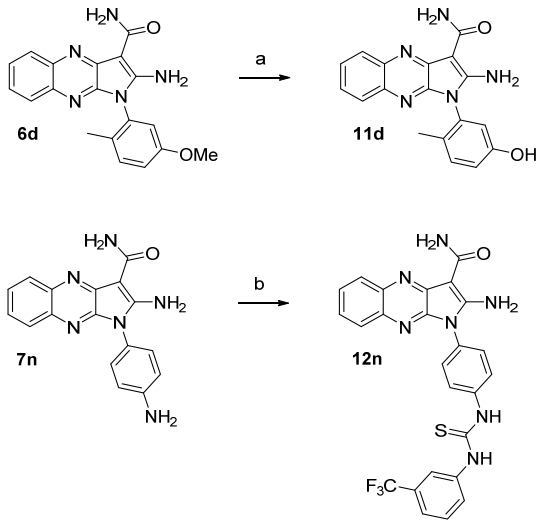



[a]Reagents and reaction conditions: Conditions a) Method I: cyclopropanecarboxylic acid, TBTU, DIPEA, DMF, 25 °C, 15h, 38%. Method II: 3-(trifluoromethyl)benzoyl chloride, $Et_3N$, DCM, 25 °C, 15 h, 54-99%. Method III: 3-(trifluoromethyl)benzoyl chloride, DIPEA, THF, 25 °C, 12 h, 48%. Method IV: benzoic acid, HOBt, EDC, DMF, 25 °C, 12 h, 64%; Conditions b) for **3a**, **3b**, **3f**, **3g**, **3k** and **3l**, $SnCl_2.H_2O$, EtOH, 100 °C, 2 h, 53-92%; for **3c-e**, **3h-j** and **3m**, 10% Pd/C (10% wt), $H_2$, MeOH, 25 °C, 4-12 h, 44-99%; Conditions c) Method I: amine, $Et_3N$, DCM, 25 °C, 15 h, 50-79%. Method II: aniline, DIPEA, THF, 25 °C, 12 h, 57-78%. Method III: (i) $SO_2Cl$, DCM, reflux, 3h; (ii) 3-(trifluoromethyl)aniline, DCM, 25 °C, 12 h, 77%; Conditions d) Method I: aniline, $Et_3N$, DCM, 25 °C, 1-3 d, 80-92%. Method II: 3-(trifluoromethyl)aniline, $Et_3N$, THF, 25 °C, 12 h, 60%.

Hydrolysis of the cyano group under strong acidic conditions furnished the desired type I and I1/2 inhibitors **6a-i** and type II inhibitors **7a-n** respectively. The presence of a cyano group in compounds **4** and **5** was confirmed by the presence of a characteristic IR band at around 2200 $cm^{-1}$ and a $^{13}$C-NMR signal at 115-116 ppm whereas the amino group appeared as a broad signal at 8 to 8.5 ppm in $^1$H-NMR.[93]

Two more inhibitors were prepared, first by demethylation of **6d** in the presence of $BBr_3$ under reflux to give the corresponding phenol derivative **11d**, and a second one by condensation of the aniline **7n** with *m*-$CF_3$-phenyl isothiocyanate, to give the corresponding thiourea **12n** (Scheme 3).

**Scheme 3a**




[a]Reagents and conditions: a) $BBr_3$, DCM, 130 °C, 4 h, 68%; b) m-CF3-phenyl isothiocyanate, DMF, 25 °C, 12 h, 69%.

### II. Biophysical characterization

The thermodynamics and kinetics of binding of the designed quinoxaline inhibitors were investigated by an array of biophysical techniques including differential scanning fluorimetry, fluorescence resonance energy transfer (FRET) based enzymatic assays and surface plasmon resonance (SPR). Differential scanning fluorimetry is a high throughput technique in which the increase in thermal stability of a folded protein upon ligand binding is detected by a fluorescence dye while measuring its melting temperature during denaturation.[97] In order to allow the binding of type II inhibitors to the inactive form of the kinase, the protein (EphA3) was incubated in the presence of the compounds for 1 hour.

The results (shown in Table 1) highlight the differences in binding between type I, $I_{1/2}$ and II compounds, with type I being the weakest binders (1.5 – 4.1 °C) and type II the most potent (with up to 16 °C thermal shifts). For type I inhibitors, the largest thermal shifts (*ca.* 4 °C) were obtained for *ortho*-methyl (**6a**) and *ortho*-chlorine (**6f**) substituted quinoxalines, which is in agreement with previously reported kinase inhibitors[26, 32, 65, 98-101] and could be a consequence of restricting the accessible conformations of the phenyl ring as we have previously reported.[32] However, the *ortho*-fluoro substituted inhibitor **6g** or bis-*ortho* substituted **6b** and **6c** caused a lower stabilization of the protein probably due to the small size of fluorine or the introduction of an extra steric bulk, respectively.

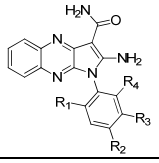
The transition from type I to type $I_{1/2}$ by the presence of a hydrogen bond donor at position $R_3$, and therefore the formation of hydrogen bonds with Glu664 and Asp758, resulted in a remarkable increase in thermal shift for compound **6h** (7.2 °C), which became

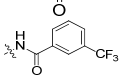
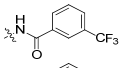
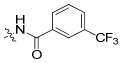
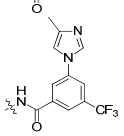
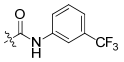
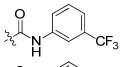
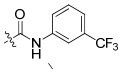
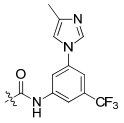
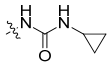
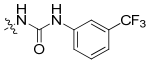
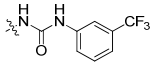
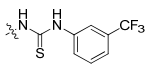
more pronounced upon introduction of a methyl group in the *ortho-* relative position (following the trend observed in type I inhibitors) to yield **11d** with a thermal shift of 11.2 °C.

Type II compounds bearing an amide linker followed by a *m*-$CF_3$-phenyl group caused a similar stabilization effect in the protein as the type $I_{1/2}$ inhibitor **11d**. As expected from previous results with type I and $I_{1/2}$ compounds, the addition of a methyl (**7c**, 7**h)** or fluorine (**7d**, 7**i**) substituent in $R_1$ lead to higher thermal shifts. Interestingly, 3-amides (**7g-i**) triggered a higher stabilization of the kinase than 1-amides (**7b-d**), which could indicate the formation of a more favorable hydrogen bond with Glu664. The presence of imidazoles within the allosteric binding site led to compounds **7e** and **7j**, which showed the most promising thermal shifts (16 °C). Urea (**7m**) or thiourea (**12n**) linkers located in *para-* relative position retained or even enhanced (in the case of urea **7m**) the binding affinity, whereas compound **7l**, bearing the urea in *meta-* relative position barely presented any thermal shift, suggesting a disruption or a non-favorable hydrogen bond interaction with Glu664. The replacement of the *m*-$CF_3$-phenyl group by a cyclopropyl ring proved to be detrimental in all cases and no thermal shift was observed for products **7a**, **7f** and **7k**.

The inhibitory activities of Type I and type $I_{1/2}$ inhibitors were further evaluated on an enzymatic assay based on fluorescence resonance transfer (FRET) at a single concentration (1 µM, Table 1, column 8). The results were in line with the thermal shifts described above. Compounds **6a** and **6f** showed inhibitory activities higher than 66% at 1 µM, (Table 1). However, substitution at $R_1$ by a smaller fluorine atom yielded compound **6g** with lower binding affinity (30%). Along the same lines, bis-*ortho* substitutions with either methyl or chlorine (**6b**, **c**) decreased the binding affinity (36 and 45% inhibitory activities respectively) probably due to unfavorable steric effects. The presence of a hydrogen bond donor at $R_3$ (type $I_{1/2}$ inhibitors) either as a phenol (**6h**, **11d**) or methylene alcohol (**6e**) greatly improved the inhibitory activity of the molecules thanks to the formation of hydrogen bonds with Glu664 and Asp758 (68-105%). To our surprise, a triazole group at the same position (**6i**) dramatically decreased the binding affinity (7%).

**Table 1.** EphA3/EphB4 inhibition data for the synthesized quinoxaline derivatives



| Cmp | Type of binding | $R_1$ | $R_2$ | $R_3$ | $R_4$ | Thermal shift (degrees)[a] | FRET[b] enzymatic assay | Cellular $IC_{50}$ (nM)[c] |
|---|---|---|---|---|---|---|---|---|
| **6a** | I | Me | H | H | H | 4.1 | 66 | 230 |
| **6b** | I | Me | H | H | Me | 2.6 | 36 | 4400 |
| **6c** | I | Me | H | H | Cl | 2.6 | 45 | 2800 |
| **6d** | I | Me | H | OMe | H | 1.5 | 23 | 720 |
| **6e** | $I_{1/2}$ | Me | H | $CH_2OH$ | H | 3.6 | 68 | 160 |
| **6f** | I | Cl | H | H | H | 3.0 | 90 | - |
| **6g** | I | F | H | H | H | 2.7 | 30 | - |
| **6h** | $I_{1/2}$ | H | H | OH | H | 7.2 | 98 | - |
| **6i** | $I_{1/2}$ | H | H |  | H | 0.1 | 7 | - |
| **11d** | $I_{1/2}$ | Me | H | OH | H | 11.2 | 105 | 5.4 |
| **7a** | II | H | H |  | H | 0.3 | - | - |
| **7b** | II | H | H |  | H | 7.6 | - | 20 |
| **7c** | II | Me | H |  | H | 12.3 | - | 14 |
| **7d** | II | F | H |  | H | 10.9 | - | 15 |
| **7e** | II | Me | H |  | H | 15.8 | - | 170 |
| **7f** | II | H | H |  | H | 0.6 | - | - |
| **7g** | II | H | H |  | H | 9.1 | - | 20 |
| **7h** | II | Me | H |  | H | 14.4 | - | 21 |
| **7i** | II | F | H |  | H | 13.0 | - | 24 |
| **7j** | II | Me | H |  | H | 16.2 | - | 270 |
| **7k** | II | H | H |  | H | 0.3 | - | - |
| **7l** | II | H | H |  | H | 0.7 | - | - |
| **7m** | II | H |  | H | H | 14.3 | - | 89 |
| **12n** | II | H |  | H | H | 9.9 | - | 560 |

[a]Average values of triplicate measurements. The standard deviation is smaller than 0.5 degrees. [b]FRET-based enzymatic assay was carried out using the Z-LYTE Kinase Assay Kit−Tyr 1 Peptide (Invitrogen) following the vendor instructions. % of inhibition at 1 µM is given. [c]Cell IC50 values were measured in a cellular phosphorylation assay using MEF cells overexpressing EphB4 at Proqinase.

The kinetics of binding of the optimized type $I_{1/2}$ (**11d**) and type II urea inhibitor **7m** were investigated using SPR. Upon titrating **11d** and **7m** over immobilized dephosphorylated EphA3, dissociation constants ($K_D$) in the low nanomolar range were obtained (8.6 nM and 39.3 nM respectively), confirming the high affinity of the compounds (Figure 15). One of the advantageous characteristics of type II inhibitors over type I or $I_{1/2}$ is their slow dissociation rate from the target protein,[24, 102, 103] as demonstrated by the remarkably slow $k_{off}$ measured for **7m** ($1.45 \times 10^{-4}$ $s^{-1}$) in comparison to the type $I_{1/2}$ **11d** ($1.52 \times 10^{-3}$ $s^{-1}$). The slow $k_{off}$ of **7m** corresponds to a residence time of 115 min, a value that compares positively with that of marketed drugs such as imatinib (28 min for dephosphorylated ABL), nilotinib (202 min and 205 min for dephosphorylated and phosphorylated ABL, respectively), and dasatinib (15 min and 4 min).[104] The long residence time of type II inhibitors is considered to be beneficial for drug efficacy and selectivity *in vivo* due to the high concentration of the drug near the target,[103, 105] as described for the EGFR-specific inhibitor lapatinib.[106]

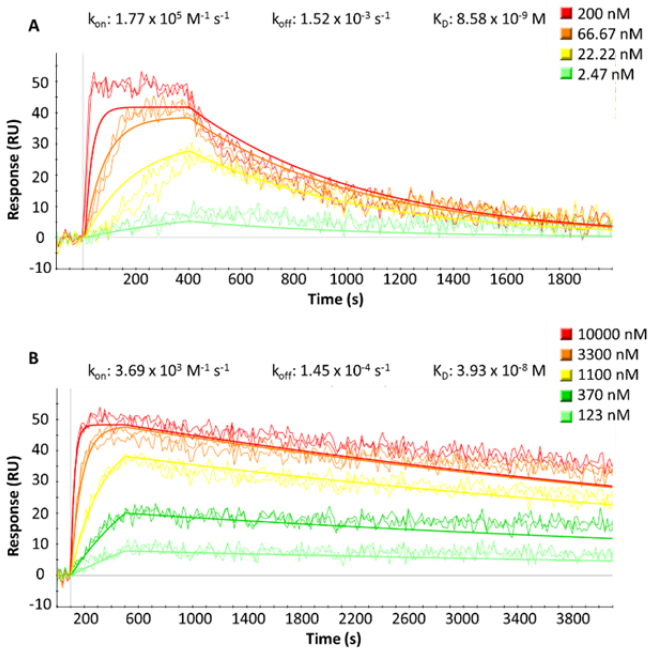



**Figure 15.** SPR analysis of the binding of **11d** (A) and **7m** (B) to EphA3. A threefold serial dilution of the compounds was made starting from 200 nM (for **11d**) and 10 uM (for **7m**) in duplicate. The theoretical global fit to a 1:1 interaction model is shown in straight lines.

### 2.2.3 Validation of type II binding by X-ray crystal structure determination

The crystal structure of the catalytic domain of the EphA3 kinase in complex with inhibitor **7m** (solved at 2.0 Å resolution), confirms a type II binding mode, i.e., with the

DFG-out conformation (Figure 16). The pyrrolo[3,2-*b*]quinoxaline scaffold occupies the ATP binding site and is involved in the same hydrogen bonds with the hinge region as the type I inhibitor **A** (Figure 14). In addition, the urea linker of inhibitor **7m** acts as hydrogen bond acceptor from the Ser757 side chain and the amide backbone of Asp758, and hydrogen bond donor to the side chain of Glu664. The *m*-$CF_3$-phenyl moiety is nestled in the hydrophobic pocket which originates from the displacement of the Phe side chain of the DFG motif. Thus, the type II binding mode of compound **7m** validates our design based on the crystal structures of the complexes with the type I inhibitors (*vide supra*, Section 2.2.1).

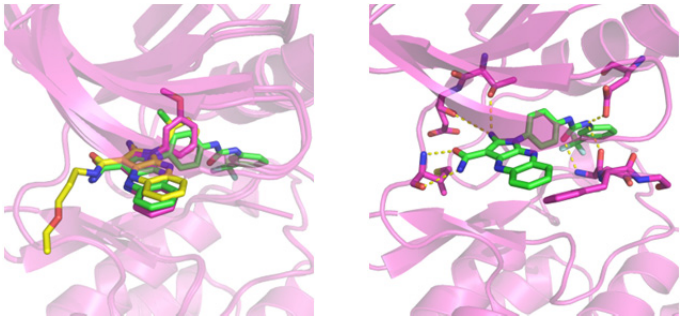


**Figure 16.** Crystal structures of the catalytic domain of the tyrosine kinase EphA3 in complex with the low-nanomolar inhibitor **7m** (left, PDB code 4P5Z) and superposition of the three inhibitors **A**, **B,** and **7m** based on structural alignment of the $C_\alpha$ atoms of the EphA3 kinase domain (right). The ATP binding site of the EphA3 kinase is shown in magenta ribbons while the side chains mentioned in the text and the inhibitors are shown by sticks

### 2.2.4 Selectivity and cellular activity

#### I. Selectivity profiles from biochemical assays

The selectivity profile of inhibitors **11d**, **7b** and **7g** was determined by an *in vitro* competition binding assay using recombinant kinases (KINOMEscan at DiscoveRx).[107] It is important to note that this assay reports on binding affinity and does not require ATP. The selectivity panel consisted of 453 human kinases, 58 of which were disease related mutant kinases (mainly of ABL1, EGFR, and PIK3CA). Single dose measurements were carried out at 1 µM concentration of the inhibitor.

Compounds **11d**, **7b** and **7g** present a very similar selectivity profile (Figure 17); in particular, strong binding is only observed to tyrosine kinases with threonine as a gatekeeper residue, e.g., ABL1/2, BRAF, DDR1, EphA/B (all but EphA7 which has a Ile gatekeeper), KIT, LCK, SRC, and YES. The latter data suggest that most (or even all) tyrosine kinases with a Thr gatekeeper can assume the DFG-out conformation. Quantitatively, each of the three inhibitors **11d**, **7b** and **7g** binds with an affinity 10-fold (100-fold) higher than the DMSO negative control to only about 10% (5%) of the 395 wild-

type kinases tested. Interestingly, the selectivity profiles of the type $I_{1/2}$ (**11d**) and II (**7b** and **7g**) quinoxaline-based inhibitors is very similar to the one of our previously reported type I and I1/2 xanthine-based inhibitors (compounds **40** and **3**, respectively, in reference 31), which is due, at least in part, to the use of an Eph tyrosine kinase (EphB4) as primary target for the *in silico* screening and optimization.

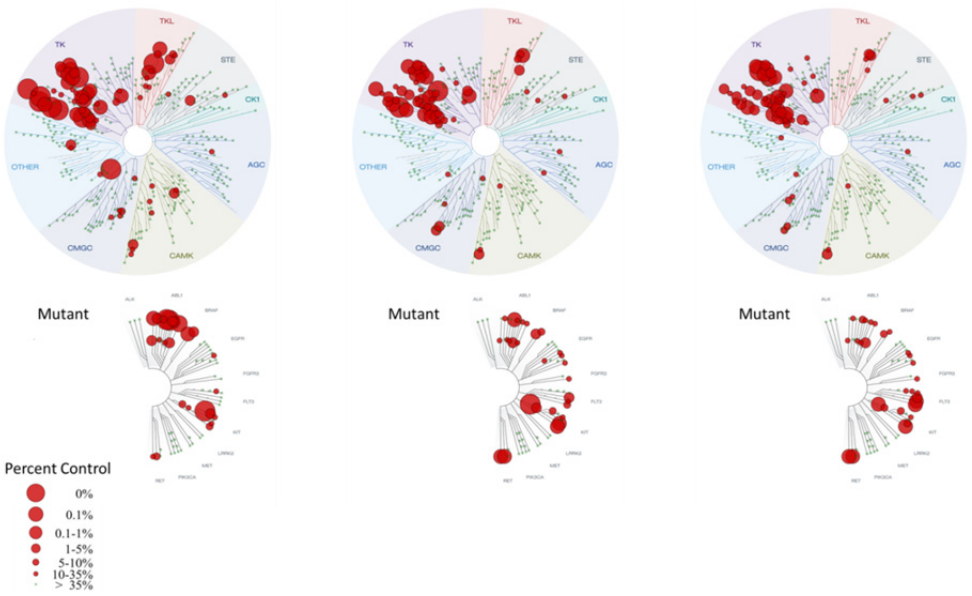



**Figure 17.** Selectivity profiles of compound **11d** (left), **7b** (center) and **7g** (right) tested on a panel of 453 protein kinases at DiscoveRx. Measurements were performed at a concentration of 1 µM of the inhibitor. The affinity is defined with respect to a DMSO control. The dendrogram was obtained from KinomeScan using the KinomeTree software.

### II. Cellular assays

The most potent inhibitors obtained in the optimization campaign were further tested in cell-based assays. Cellular phosphorylation assays on MEF cells transfected with myc-tagged human EphB4 revealed a comparable tendency to the one observed in the enzymatic assay (Table 1, column 9). The type I inhibitors (**6a-d**) displayed cellular $IC_{50}$ values in the 230-4400 nM range, with the *ortho* methyl substituted derivative **6a** as the most potent member of this series. The type $I_{1/2}$ inhibitor **11d** and type II compounds bearing amide linkers and a *m*-$CF_3$-phenyl group (**7b-d**, **7g-i**) displayed levels of inhibitory activity in the low nanomolar range (6-24 nM), thus being the most promising molecules of the optimization campaign. In agreement with thermal shift experiments, the presence of an urea (**7m**) and specially thiourea linker (**12n**) decreased the potency of the compounds (89 and 560 nM respectively). Interestingly, the imidazole substituted compounds (**7e**, **7j**) proved to be the weakest type II inhibitors (170 and 270 nM respectively) in contrast to the high thermal shifts obtained (16 °C) in the differential scanning fluorimetry measurements, pointing towards potential cell permeability or efflux issues.

EphB4 overexpression has been linked to several types of cancer, including breast,[49] colon[108] and ovarian.[109] Compounds **11d** and **7m** were screened against the NCI-60 cancer cell line panel[110] displaying antiproliferative activities against leukemia (K-562), lung (HOP-92), colon (HT-29), renal (A498) and breast cancer cells (MDA-MB-231, HS 578T) in the low nM range.

**Table 2**. Antiproliferative activity against NCI tumor cell lines[a]

| Cmp | MDA-MB-231 | K-562 | A498 | HT29 | KM12 | HOP-92 |
|---|---|---|---|---|---|---|
| **6d** | | | 59.3 | 50.3 | | |
| **6e** | | 44.9 | 36.4 | 90.5 | | |
| **6i** | | | 50.9 | 64.5 | | |
| **11d** | 2.64 | 1.05 | 5.88 | 4.59 | 1.55 | 0.49 |
| **7b** | 3.09 | 1.52 | 13.4 | 29.4 | 3.98 (4.03) | 4.65 (3.91) |
| **7c** | 3.88 | 0.73 | 2.50 | 5.87 | 0.80 | 4.23 |
| **7d** | 1.44 | 0.37 | 1.88 | 1.80 (9.43) | 2.16 | 2.06 |
| **7e** | 1.93 | 0.030 | 4.01 | 13.4 | 2.54 | 5.12 |
| **7g** | 10.8 | 0.820 | 10.6 | 23.1 | 2.78 | 5.13 |
| **7h** | 1.32 | 0.036 | 2.07 | 2.82 | 1.57 | 1.92 |
| **7i** | 2.69 | 0.081 | 2.44 | 2.97 | 1.87 | 2.12 |
| **7j** | 3.05 | 0.029 | 5.71 | 4.36 | 2.84 | 3.01 |
| **7m** | 10.3 | 5.45 | 10.8 | 18.0 | 0.67 | 10.0 |
| **12n** | 9.63 | 5.07 | 8.73 | 16.3 | 2.52 | 14.25 |

[a] GI50 values were determined using resazurin reduction after 2-3 days of incubation with the corresponding compound. GI50 values are given in micromolar concentrations (μM) as the mean of at least three independent experiments. Variability around the mean value was <50% unless otherwise indicated by an SE value in parentheses.

Driven by these results, the most promising inhibitors of our optimization campaign were tested in-house against the above mentioned NCI cancer cell lines (Table 2). The leukemia K-562 cell line was particularly sensitive towards the optimized type II quinoxaline inhibitors, especially in the case of 3-amide compounds **7h** and **7i**, which showed remarkably low $GI_{50}$ values (36 and 81 nM respectively). Interestingly, similar levels of potency were found for imidazole substituted compounds **7e** and **7j**, which seemed to be among the weakest type II binders in the cellular phosphorylation assays, possibly indicating other targets than Eph for these molecules.

In addition, the potential of **11d** (the most potent compound on cellular phosphorylation assays with an $IC_{50}$ of 5.4 nM) to inhibit the growth of patient-derived

tumor cell lines was studied using a propidium iodide-based proliferation assay and dasatinib as a reference (Oncotest, Table 3). Cell lines included colon, lung, kidney, pancreatic, prostate, and stomach cancer cells. Whereas dasatinib presented double-digit nanomolar activities against RXF 393NL, LXFA 983L, and PRXF DU145, **11d** exhibited low micromolar $GI_{50}$ values, with RXF 393NL being the most sensitive cell line.

**Table 3.** Antiproliferative activity against patient derived tumor cell lines[a]

| Compound | **11d** | **dasatinib** |
|---|---|---|
| RXF 393NL | 0.72 | 0.0217 |
| CXF 1103L | 3.83 | 4.36 |
| LXFA 983L | 2.22 | 0.0565 |
| GXF 251L | 8.01 | 2.25 |
| PAXF 1657L | 2.92 | 0.121 |
| PRXF DU145 | 2.92 | 0.0623 |

[a]$GI_{50}$ values were determined at Oncotest using a modified propidium iodide assay. Measurements were performed after 4 days of incubation with **11d** and dasatinib. $GI_{50}$ values are given in micromolar concentrations (µM).

The implication of EphB4-ephrinB2 signaling in sprouting angiogenesis and blood vessel maturation[111] and the inhibition of vascular endothelial growth factor (VEGFR)-driven angiogenesis by the selective EphB4 inhibitor NVP-BHG712,[64] led us to examine the efficacy of **11d** on human endothelial cell sprouting in a spheroid based cellular angiogenesis assay (ProQinase).[112] Compound **11d** was able to successfully inhibit VEGF-A induced HUVEC (primary human umbilical vein endothelial cells) sprouting in a dose dependent manner with an $IC_{50}$ value of 1.5 µM.

### 2.2.5 In vivo data

Three of the most promising compounds from these series (**7b**, **7g** and **11d**) were selected for evaluation of pharmacokinetic properties in 20-30 g male CD-1 (ICR) mice on intravenous (IV) and oral (PO) administration. Low to moderate oral bioavailability of tested compounds in mice was observed, with compounds **11d** and **7g** giving the highest values (Table 4). Promising cellular efficacy and pharmacokinetic properties incited the subsequent evaluation of compound **11d** in a xenograft mouse model with a tumor derived from the MDA-MB-231 cell line. High compound clearance (Cl) and moderate half-life ($T_{1/2}$) values determined in the pharmacokinetic study motivated a twice-daily dosing regime totaling 100 mg/kg/day of compound **11d** over 21 days. Median tumor volume progression over time, starting from 108 $mm^3$, of both treatment and control cohorts is given in Figure 18.

**Table 4**. Pharmacokinetic properties in mice

| Compound | **11d** | | **7b** | | **7g** | |
|---|---|---|---|---|---|---|
| | IV | PO | IV | PO | IV | PO |
| Dose (mg/kg) | 1 | 5 | 1 | 5 | 1 | 5 |
| Cl (mL/min/kg) | 42 | - | 32 | - | 31 | - |
| $V_{ss}$ (L/kg) | 1.6 | - | 2.2 | - | 2.2 | - |
| $T_{1/2}$ (h) | 1.7 | 1.7 | 1.2 | 5.0 | 1.1 | 2.8 |
| AUClast (h.ng/mL) | 392 | 493 | 506 | 263 | 533 | 803 |
| F(%) | - | 25 | - | 10 | - | 30 |

In this study, median treatment-group tumor volume remained essentially stable throughout the treatment period, achieving a median tumor volume of 126 $mm^3$ at day 19 as opposed to the control group whose median tumor reached 650 $mm^3$ in the same period. Tumor growth inhibition (%TGI) was statistically significant (Mann-Whitney U=0, $P \leq 0.001$ two-tailed) and quantified at 81% relative to the control group. Mean body weight of the treatment cohort decreased up to 16.3% of the initial mean body weight of this cohort during the treatment period. Treatment with compound **11d** provides a significant limitation in tumor progression over the control, suggesting that further studies of such xenograft model at lower doses of compound **11d** might provide tumor volume control with lessened weight loss. The further evaluation of compounds **11d** on mouse models of K-562 leukemia is underway.

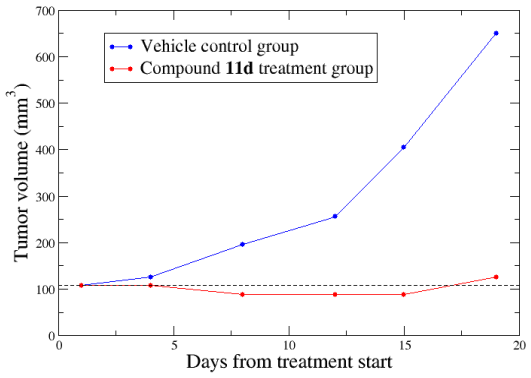



**Figure 18.** *In vivo* antitumor activity of compound **11d** in MDA-MB-231 nude mice xenografts. The mice received by gavage twice-daily 50 mg/kg of compound **11d** (red) or vehicle control (blue). Each data point is the median of a cohort of 9 animals.

### 2.2.6 Conclusions

The X-ray crystal structures of the EphA3 kinase in complex with two high-nanomolar inhibitors based on the 2-amino-1-phenyl-pyrrolo[3,2-*b*]quinoxaline-3-carboxamide scaffold confirmed the type I binding mode obtained previously by automatic docking (Figure 14). This structural information was used to design type $I_{1/2}$ and type II derivatives by taking advantage of the existing knowledge on privileged chemical motifs, i.e., hydroxyl group in *meta* position of the phenyl ring (for type $I_{1/2}$) and hydrophobic moieties connected to the phenyl ring by amide or urea linkers (on type II). Chemical synthesis of *ca.* 25 derivatives (Table 1) culminated in several low nanomolar inhibitors with a good selectivity profile (Figure 17). The X-ray crystal structure of the EphA3 kinase in the complex with the inhibitor **7m** (Figure 16) provided the final validation of the structure-based design, in particular the DFG-out conformation confirmed the type II binding. Moreover, the slow kinetics of unbinding of compound **7m** (measured by SPR, Figure 15) is congruent with the type II binding mode. Three interesting observations emerge from this study. First, it is possible to "elongate" a type $I_{1/2}$ into a type II inhibitor by introducing an amide or urea linked to a bulky hydrophobic group. These type II linkers are involved in the same hydrogen bonds as the type $I_{1/2}$ bearing a hydroxyl group in the same position, while the hydrophobic moiety occupies the pocket resulting from the displacement of the Phe side chain of the DFG motif. The similar selectivity profiles of type $I_{1/2}$ and type II inhibitors indicate that mainly the moiety in contact with the gatekeeper's side chain and hinge region determines specificity. Finally, *in vivo* assays (mice xenografted with human breast cancer) confirmed the cytostatic activity of one of our inhibitors (**11d**) which makes this type $I_{1/2}$ compound a candidate lead for further preclinical development.

## 2.3 Experimental section

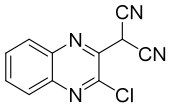
### 2.3.1 Chemistry. General methods.

All reactions, unless otherwise stated, were carried out under a nitrogen atmosphere using standard Schlenk-techniques. All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 $F_{254}$. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV 300, AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in

ppm. The spectra are calibrated to the residual $^{1}H$ and $^{13}C$ signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Buchi melting point B-540 instrument. High-resolution electrospray ionization mass spectrometry was performed on a Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. Ten spectra were acquired. A mass accuracy ≤2 ppm was obtained in the peak matching acquisition mode by using a solution containing 2 µL PEG200, 2 µL PPG450, and 1.5 mg NaOAc (all obtained from Sigma-Aldrich, Buchs, Switzerland) dissolved in 100 mL MeOH (HPLC Supra grade, Scharlau, E-Barcelona) as internal standard. The purity of all tested compounds was determined by HPLC on a Waters Acquity UPLC (Waters, Milford, MA) Top spectrometer using an Acquity BEH C18 HPLC column (1.7 µm, 1× 50 mm, Waters) with a mixture of $H_2O$ + 0.1% HCOOH (A) and $CH_3CN$ + 0.1% HCOOH (B) solvent (0.1 mL flow rate, linear gradient from 5% to 98% B within 4 min followed by flushing with 98% B for 1 min). Unless otherwise stated, all compounds showed ≥ 95 % purity.

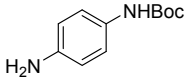
The following compounds were prepared according to previously reported procedures: **1**[92, 113], **8e**[95], **13**,[114] **9e**,[114] **3n**,[115] **14c**,[116] **14d**,[117] **14h**,[118] **14m**,[119] **3c**,[114] **3h**,[118] **3m**.[119]

**2-(3-Chloroquinoxalin-2(1*H*)-ylidene)malononitrile (1)[92, 113]**



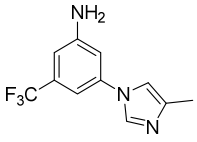

Yellow solid; Yield: 74 %; $^{1}H$ NMR (400 MHz, DMSO-$d_6$): δ = 7.67-7.62 (m, 2H), 7.57 (ddd, $J$ = 8.4, 6.8, 1.1 Hz, 1H), 7.36 (ddd, $J$ = 8.1, 6.8, 1.2 Hz, 1H), NH-not observed; $^{13}C$ NMR (100 MHz, DMSO-$d_6$): δ = 149.3, 141.9, 134.8, 134.1, 131.3, 127.5, 125.5, 120.2, 118.6, 44.5; IR (film): $\tilde{\upsilon}$ = 3235, 3203, 3100, 3056, 3013, 2979, 2220, 2208, 1616, 1577, 1488, 1410, 1096, 969, 799, 764, 594 $cm^{-1}$; MS (ESI): *m/z*: calcd for $C_{11}ClH_5N_4Na^+$: 251.0, found: 250.9.

***N*-Boc-*p*-phenylenediamine (3n)[115]**



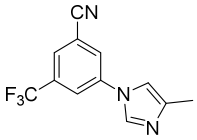

Yellow solid; Yield: 97 %; $^{1}H$ NMR (500 MHz, DMSO-$d_6$): δ = 8.76 (br, 1H), 7.06 (d, $J$ = 6.7 Hz, 2H), 6.46 (d, $J$ = 8.8 Hz, 2H), 4.71 (br, 2H), 1.44 (s, 9H); $^{13}C$ NMR (125 MHz, DMSO-$d_6$): δ = 153.1, 143.9, 128.6, 120.2, 114.0, 78.1, 28.2; IR (neat): $\tilde{\upsilon}$ = 3416, 3394, 3364, 3188, 2984, 2935, 1692, 1521, 1512, 1427, 1391, 1367, 1307, 1233, 1158, 1055, 1028, 1013, 905, 823, 772, 761, 727, 615, 511, 417 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{11}H_{16}N_2NaO_2^+$: 231.11, found: 231.0.

**3-(4-Methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)phenylamine (8e)[95]**



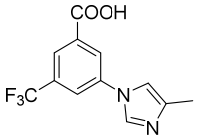
Yellow solid; Yield: 73 %; $^{1}$H NMR (300 MHz, DMSO-$d_6$): δ = 8.06 (s, 1H), 7.36 (s, 1H), 6.97 (s, 1H), 6.93 (s, 1H), 6.81 (s, 1H), 5.85 (s, 2H), 2.15 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 150.9, 138.5, 134.7, 131.2 (q, $J$ = 31.5 Hz), 124.0 (q, $J$ = 272.5 Hz), 114.2, 107.8, 103.2 (q, $J$ = 3.7 Hz), 13.6, 2C are missing due to overlapping; $^{19}$F NMR (376 MHz, DMSO-$d_6$): δ = -61.63; IR (neat): ϑ = 3306, 3159, 2981, 1696, 1575, 1540, 1515, 1445, 1417, 1330, 1285, 1245, 1222, 1162, 1114, 1091, 1054, 902, 849, 808, 710, 695, 663, 602, 468, 456, 421, 408 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{11}H_{10}F_3N_3$: 242.1, found: 242.2.

**3-(4-Methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)benzonitrile (13)[114]**



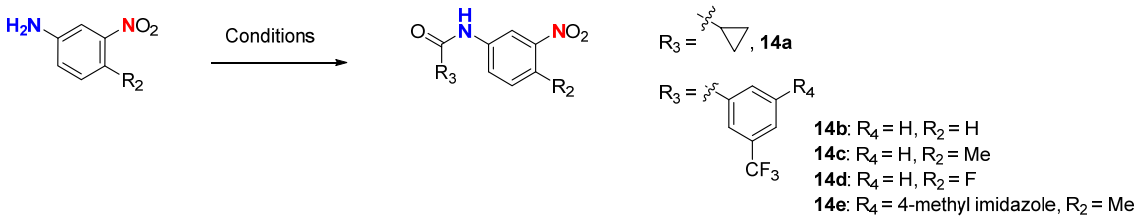
White solid; Yield: 71 %; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 8.54 (s, 1H), 8.41 (d, $J$ = 1.4 Hz, 1H), 8.37 (s, 1H), 8.24 (s, 1H), 7.71 (t, $J$ =1.2 Hz, 1H), 2.16 (d, $J$ = 0.9 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO): δ = 139.1, 138.2, 135.3, 131.8 (q, $J$ = 33.5 Hz), 126.8, 126.5 (q, $J$ = 3.8 Hz), 122.8 (q, $J$ = 273.3 Hz), 120.6 (q, $J$ = 3.6 Hz), 117.0, 114.1, 114.0, 13.6; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.34; IR (neat): ϑ = 3107, 3045, 2927, 2242, 1614, 1606, 1504, 1464, 1402, 1319, 1290, 1267, 1228, 1171, 1134, 1112, 1077, 1012, 880, 835, 805, 749, 705, 691, 626, 595, 517, 460, 411 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{12}H_9F_3N_3^+$: 252.1, found: 251.9.

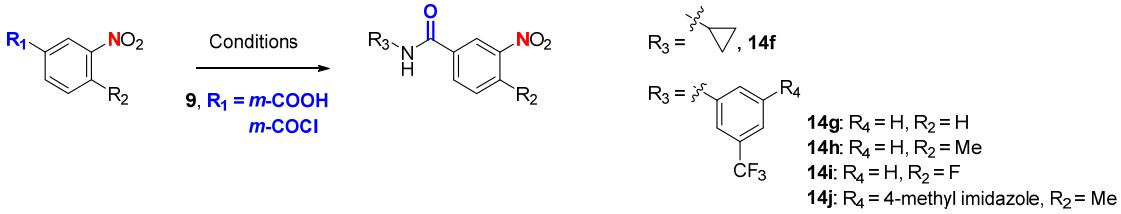
**3-(4-Methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)benzoic acid (9e)[114]**



White solid; Yield: 93 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 14.03 (br, 1H), 9.75 (s, 1H), 8.56 (s, 1H), 8.50 (s, 1H), 8.29 (s, 1H), 8.21 (s, 1H), 2.36 (d, $J$ = 0.8 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.9, 136.3, 134.6, 134.1, 131.0 (q, $J$ = 33.4 Hz), 130.7, 126.4, 125.9 (q, $J$ = 3.2 Hz), 123.2 (q, $J$ = 3.2 Hz), 123.1 (q, $J$ = 273.3 Hz), 117.5, 9.8; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.30; IR (neat): ϑ = 3096, 2769, 2631, 1708, 1627, 1616, 1547, 1404, 1348, 1304, 1270, 1199, 1169, 1127, 1085, 998, 962, 910, 844, 769, 709, 691, 673, 659, 620, 527 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{12}H_8F_3N_2O_2$: 269.1, found: 268.9.

## I. Synthesis of Amides

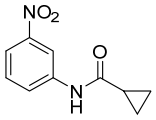



### General procedure for the synthesis of amides

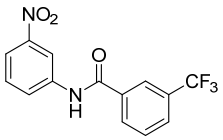
To a solution of primary amine (1.00 eq) in $CH_2Cl_2$ (0.500 M) triethylamine (1.05 eq) and acid chloride (1.04 eq) were sequentially added. The reaction was stirred at 25 °C for 15 h. The reaction mixture was then poured into water, extracted with $CH_2Cl_2$, and the organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The residue was recrystallized from hexane/EtOAc to afford the desired products in pure form. This method was used to obtain intermediates **14b**, **14g** and **14f**.

### *N*-(3-Nitrophenyl)cyclopropanecarboxamide (14a)



To a suspension of cyclopropanecarboxylic acid (100 µL, 1.26 mmol) and TBTU (556 mg, 1.73 mmol) in DMF (5.00 mL) 3-nitroaniline (174 mg, 1.26 mmol) and *N,N*-diisopropylethylamine (2.10 mL, 12.1 mmol) were added. The reaction mixture was stirred at 25 °C for 15 h, poured into a solution of saturated $NaHCO_3$ and extracted with EtOAc. The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography (hexane/EtOAc 8:1) afforded the desired compound in pure form as a white solid (200 mg, 38 %). $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.68 (s, 1H), 8.63 (t, $J$ = 2.2 Hz, 1H), 7.91-7.87 (m, 2H), 7.59 (t, $J$ = 8.2 Hz, 1H), 1.79 (quint, $J$ = 6.2 Hz, 1H), 0.86-0.85 (m, 4H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 172.3, 147.9, 140.3, 130.0, 124.7, 117.3, 112.9, 14.6, 7.5; IR (film): $\tilde{\upsilon}$ = 3284, 1655, 1521, 1317, 1194, 960, 673 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{10}H_{11}N_2O_3^+$: 207.2, found: 206.9.

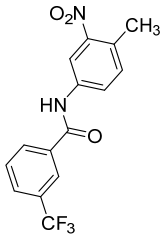
### *N*-(3-Nitrophenyl)-3-(trifluoromethyl)benzamide (14b)



White solid; Yield: 54 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.88 (s, 1H), 8.77 (t, $J$ = 2.2 Hz, 1H), 8.34 (s, 1H), 8.30 (d, $J$ = 7.8 Hz, 1H),

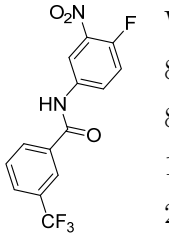
8.22-8.20 (m, 1H), 8.02-7.99 (m, 2H), 7.82 (t, $J$ = 7.8 Hz, 1H), 7.69 (t, $J$ = 8.2 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.4, 147.8, 139.9, 135.0, 131.9, 130.1, 129.8, 129.2 (q, $J$ = 32.1 Hz), 128.5 (q, $J$ = 3.5 Hz), 126.2, 124.2 (q, $J$ = 3.8 Hz), 123.8 (q, $J$ = 272.6 Hz), 118.4, 114.5; IR (film): $\tilde{\upsilon}$ = 3292, 1653, 1524, 1337, 1264, 1122, 1072 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{14}H_9F_3N_2NaO_3^+$: 333.2, found: 333.0.

**_N_-(3-Nitro-4-methylphenyl)-3-(trifluoromethyl)benzamide (14c)[116]**



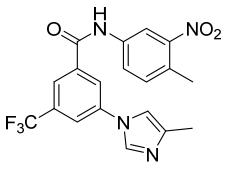
Pale yellow solid; Yield: 99 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.80 (s, 1H), 8.53 (d, $J$ = 2.1 Hz, 1H), 8.33 (s, 1H), 8.29 (d, $J$ = 7.9 Hz, 1H), 8.03-7.99 (m, 2H), 7.81 (t, $J$ = 7.8 Hz, 1H), 7.52 (d, $J$ = 8.4 Hz, 1H), 2.51 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.4, 148.5, 137.9, 135.2, 133.1, 132.0, 129.9, 129.4 (q, $J$ = 32.2 Hz), 128.5 (d, $J$ = 3.3 Hz), 128.1, 125.1, 124.4 (q, $J$ = 3.7 Hz), 124.0 (q, $J$ = 272.5 Hz), 115.8, 19.4; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.08; IR (neat): $\tilde{\upsilon}$ = 3325, 1656, 1617, 1587, 1524, 1443, 1386, 1338, 1319, 1308, 1260, 1238, 1160, 1124, 1111, 1069, 907, 889, 832, 821, 796, 751, 697, 670, 610, 561, 548, 534, 411 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{15}H_{10}F_3N_2O_3^-$, $[M-H]^{-1}$: 323.1, found: 323.0.

**_N_-(3-Nitro-4-fluorophenyl)-3-(trifluoromethyl)benzamide (14d)[117]**



White solid; Yield: 48 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.85 (s, 1H), 8.66 (dd, $J$ = 6.9, 2.6 Hz, 1H), 8.32 (s, 1H), 8.28 (d, $J$ = 7.8 Hz, 1H), 8.17-8.14 (m, 1H), 8.00 (d, $J$ = 7.8 Hz, 1H), 7.81 (t, $J$ = 7.8 Hz, 1H), 7.62 (dd, $J$ = 10.8, 9.4 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.4, 150.8 (d, $J$ = 259.1 Hz), 136.3 (d, $J$ = 7.9 Hz), 135.6 (d, $J$ = 3.2 Hz), 134.9, 131.9, 129.9, 129.3 (q, $J$ = 32.2 Hz), 128.6 (q, $J$ = 3.5 Hz), 127.8 (d, $J$ = 8.3 Hz), 124.3 (q, $J$ =3.8 Hz), 123.9 (q, $J$ = 272.5 Hz), 118.8 (d, $J$ = 21.9 Hz), 117.0 (d, $J$ = 2.9 Hz); $^{19}$F NMR (282 MHz, DMSO): δ = -61.09, -124.31 (ddd, $J$ = 11.1, 6.9, 4.0 Hz); IR (neat): $\tilde{\upsilon}$ = 3266, 3145, 3078, 1655, 1616, 1532, 1406, 1346, 1325, 1255, 1170, 1121, 1070, 919, 889, 825, 814, 696, 665, 545, 497, 415 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_7F_4N_2O_3^-$, $[M-H]^{-1}$: 327.0; found: 326.9.

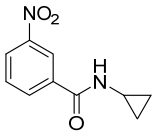
**_N_-(3-Nitro-4-methylphenyl)-3-(4-methyl-1_H_-imidazol-1-yl)-5-(trifluoromethyl)benzamide (14e)**



To a solution of 3-(4-methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)benzoic acid (**9e**, 1.06 g, 3.91 mmol), HOBt (0.634 g, 4.69 mmol) and EDC (0.728 mL, 4.69 mmol) in DMF (18.0 mL), 4-methyl-3-nitroaniline (0.714 g, 4.69 mmol) was added and the mixture was stirred at 25 °C for 12 hours. The reaction mixture was concentrated under reduced pressure, redissolved in EtOAc and extracted with 0.5 M $NaHCO_3$ and water. The organic phase was dried over $MgSO_4$, concentrated under reduced pressure and the resulting solid was washed with DCM and hexane affording the desired amide in pure form as a yellow solid (1.01 g, 64 % yield). $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.79 (s, 1H), 8.50 (d, $J$ = 1.7 Hz, 1H), 8.45 (s,

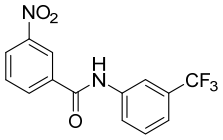
1H), 8.39 (s, 1H), 8.25 (s, 1H), 8.18 (s, 1H), 8.00 (dd, $J$ = 8.3, 1.7 Hz, 1H), 7.69 (s, 1H), 7.52 (d, $J$ = 8.4 Hz, 1H), 2.51 (s, 3H), 2.19 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.5, 148.5, 139.0, 137.8, 137.5, 136.9, 135.2, 133.2, 130.9 (q, $J$ = 32.6 Hz), 128.3, 125.0, 123.4 (q, $J$ = 273.4 Hz), 122.7, 121.8 (q, $J$ = 3.7 Hz), 119.4 (q, $J$ = 3.4 Hz), 115.7, 114.2, 19.3, 13.5; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.05; IR (neat): ῡ = 3344, 1682, 1654, 1603, 1537, 1524, 1498, 1448, 1407, 1305, 1283, 1255, 1172, 1121, 1103, 1079, 1001, 904, 877, 829, 812, 751, 739, 690, 671, 617, 503, 418 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{16}F_3N_4O_3^+$: 405.1169, found: 405.1172.

**_N_-Cyclopropyl-3-nitrobenzamide (14f)**



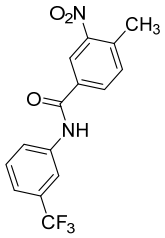
White solid; Yield: 50 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.81 (s, 1H), 8.65-8.64 (m, 1H), 8.38-8.36 (m, 1H), 8.26 (dd, $J$ = 7.9, 1.0 Hz, 1H), 7.76 (t, $J$ = 7.9 Hz, 1H), 2.92-2.86 (m, 1H), 0.75-0.71 (m, 2H), 0.63-0.60 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 165.2, 147.6, 135.7, 133.5, 129.9, 125.7, 121.7, 23.1, 5.6; IR (film): ῡ = 3227, 1626, 1518, 1347, 1315, 668 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{10}H_{10}N_2NaO_3^+$: 229.2, found: 228.9.

**3-Nitro-_N_-(3-(trifluoromethyl)phenyl)benzamide (14g)**



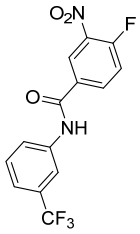
White solid; Yield: 79 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.86 (s, 1H), 8.83 (t, $J$ = 1.9 Hz, 1H), 8.48-8.42 (m, 2H), 8.24 (s, 1H), 8.07 (d, $J$ = 8.3 Hz, 1H), 7.87 (t, $J$ = 8.0 Hz, 1H), 7.64 (t, $J$ = 8.0 Hz, 1H), 7.50 (d, $J$ = 7.8 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.6, 147.7, 139.4, 135.7, 134.2, 130.2, 129.9, 129.3 (q, $J$ = 31.6 Hz), 126.4, 124.0 (q, $J$ = 272.3 Hz), 123.9, 122.4, 120.3 (q, $J$ = 3.8 Hz), 116.5 (q, $J$ = 4.1 Hz); IR (film): ῡ = 3278, 3081, 1651, 1529, 1324, 1176, 1114, 694 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{14}H_9F_3N_2NaO_3^+$: 333.2, found: 333.0.

**3-Nitro-4-methyl-_N_-(3-(trifluoromethyl)phenyl) benzamide (14h)[118]**



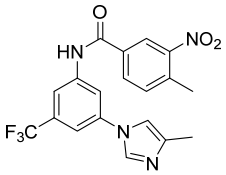
White solid; Yield: 78 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.75 (s, 1H), 8.60 (s, 1H), 8.22 (s, 2H), 8.06 (d, $J$ = 7.6 Hz, 1H), 7.69 (d, $J$ = 8.0 Hz, 1H), 7.61 (t, $J$ = 7.9 Hz, 1H), 7.47 (d, $J$ = 7.6 Hz, 1H), 2.60 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.6, 148.8, 139.6, 136.7, 133.3, 133.2, 132.2, 130.0, 129.4 (q, $J$ = 31.6 Hz), 124.1 (q, $J$ = 272.7 Hz), 124.0, 123.6, 120.3 (q, $J$ = 3.7 Hz), 116.6 (q, $J$ = 3.9 Hz), 19.6; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.27; IR (neat): ῡ = 3303, 3078, 1656, 1616, 1567, 1530, 1494, 1348, 1321, 1270, 1229, 1166, 1152, 1116, 1096, 1069, 898, 882, 861, 842, 795, 740, 696, 661, 646, 632, 524 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{15}H_{10}F_3N_2O_3^-$, $[M-H]^{-1}$: 323.1, found: 323.0.

**3-Nitro-4-fluoro-*N*-(3-(trifluoromethyl)phenyl) benzamide (14i)**



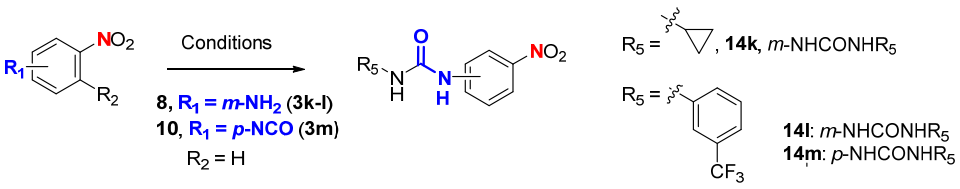
To a solution of 4-fluoro-3-nitrobenzoic acid (1.00 g, 5.40 mmol) in DCM (2.50 mL) thionyl chloride (6.42 g, 54.0 mmol) was added and it was refluxed for 3 h. The reaction was concentrated under reduced pressure and a solution containing 3-(trifluoromethyl)aniline (0.870 g, 5.40 mmol) and $Et_3N$ (1.13 mL, 8.10 mmol) in DCM (5.00 mL) was added. The reaction mixture was stirred at 25 °C for 12 hours and poured into water. The resulting solid was filtered off and washed with water, $Et_2O$ and hexane affording 3-nitro-4-fluoro-*N*-(3-(trifluoromethyl)phenyl) benzamide **14i** in pure form as a yellow solid (1.37 g, 77 % yield). $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.82 (s, 1H), 8.79 (dd, $J$ = 7.2, 2.3 Hz, 1H), 8.42 (ddd, $J$ = 8.6, 4.2, 2.3 Hz, 1H), 8.21 (s, 1H), 8.06 (d, $J$ = 8.2 Hz, 1H), 7.80 (dd, $J$ = 11.0, 8.8 Hz, 1H), 7.63 (t, $J$ = 8.1 Hz, 1H), 7.50 (d, $J$ = 8.0 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 162.9, 156.7 (d, $J$ = 267.1 Hz), 139.5, 136.7 (d, $J$ = 7.9 Hz), 135.9 (d, $J$ = 10.1 Hz), 131.2 (d, $J$ = 3.7 Hz), 130.1, 129.6 (q, $J$ = 31.3 Hz), 125.9, 124.2 (q, $J$ = 272.3 Hz), 124.1, 120.6 (q, $J$ = 3.6 Hz), 119.1 (d, $J$ = 21.5 Hz), 116.7 (q, $J$ = 4.1 Hz); $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.29, -114.01 (ddd, $J$ = 11.3, 7.3, 4.3 Hz); IR (neat): ῡ = 3281, 3083, 1653, 1618, 1536, 1491, 1445, 1326, 1257, 1167, 1117, 1094, 1069, 899, 846, 799, 749, 672, 659, 628, 528 cm$^{-1}$. HRMS (ESI), *m/z*: calcd for $C_{14}H_7F_4N_2O_3$: 327.0398, found: 327.0401.

**3-Nitro-4-methyl-*N*-(3-(4-methyl-1*H*-imidazol-1-yl)-5(trifluoromethyl))benzamide (14j)**



To a solution of 3-(4-methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)phenylamine (**8e**, 0.982 g, 4.07 mmol) in THF (14.0 mL) at 0 °C DIPEA (1.38 mL, 8.14 mmol) and 4-methyl-3-nitrobenzoyl chloride (0.711 mL, 4.88 mmol) were added. The reaction was stirred at 25 °C for 12 h and poured into water. The resulting solid was filtered off and washed with DCM and 1M HCl affording the corresponding amide in pure form as a white solid (0.940 g, 57 % yield). $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 11.38 (s, 1H), 9.65 (s, 1H), 8.68 (s, 1H), 8.64 (s, 1H), 8.36 (br, 2H), 8.02 (s, 1H), 7.96 (s, 1H), 7.72 (d, $J$ = 8.0 Hz, 1H), 2.61 (s, 3H), 2.37 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.9, 148.8, 141.2, 137.1, 136.0, 134.3, 133.3, 132.6, 132.4, 131.2, 130.8 (q, $J$ = 32.8 Hz), 125.0, 123.9, 123.4 (q, $J$ = 272.8 Hz), 117.4, 117.3, 114.2 (q, $J$ = 3.4 Hz), 19.6, 9.8; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.44; IR (neat): ῡ = 2991, 2762, 2620, 1675, 1612, 1569, 1529, 1473, 1444, 1380, 1346, 1299, 1245, 1230, 1167, 1121, 1093, 929, 887, 837, 797, 737, 715, 697, 666, 639, 583, 460, 406 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{16}F_3N_4O_3^+$: 405.1169, found: 405.1169.

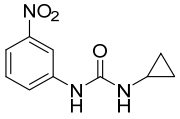
### II. Synthesis of Ureas




**General procedure for the synthesis of ureas**

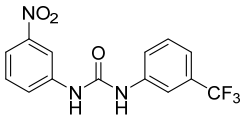
To a solution of aniline (1 eq) in $CH_2Cl_2$ (0.3 M) the isocyanate (1 eq) was added. The reaction mixture was stirred at 25 °C for 1-3 days, and the resulting precipitate was filtered off affording the desired compound in pure form. This method was used to obtain intermediates **14k** and **14l**.

**1-Cyclopropyl-3-(3-nitrophenyl)urea (14k)**



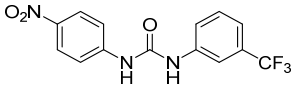

White solid; Yield: 80 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.84 (s, 1H), 8.51 (t, $J$ = 2.2 Hz, 1H), 7.75 (dd, $J$ = 8.2, 2.2 Hz, 1H), 7.69 (d, $J$ = 8.2 Hz, 1H), 7.50 (t, $J$ = 8.2 Hz, 1H), 6.59 (s, 1H), 2.58-2.54 (m, 1H), 0.67-0.63 (m, 2H), 0.45-0.42 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 155.7, 148.0, 141.6, 129.7, 123.8, 115.5, 111.7, 22.3, 6.3; IR (film): ῡ = 3348, 3264, 1648, 1515, 1344, 1243, 731 cm$^{-1}$; MS (ESI): $m/z$: calcd for $C_{10}H_{11}N_3NaO_3^+$: 244.2, found: 243.9.

**1-(3-Nitrophenyl)-3-(3-(trifluoromethyl)phenyl)urea (14l)**



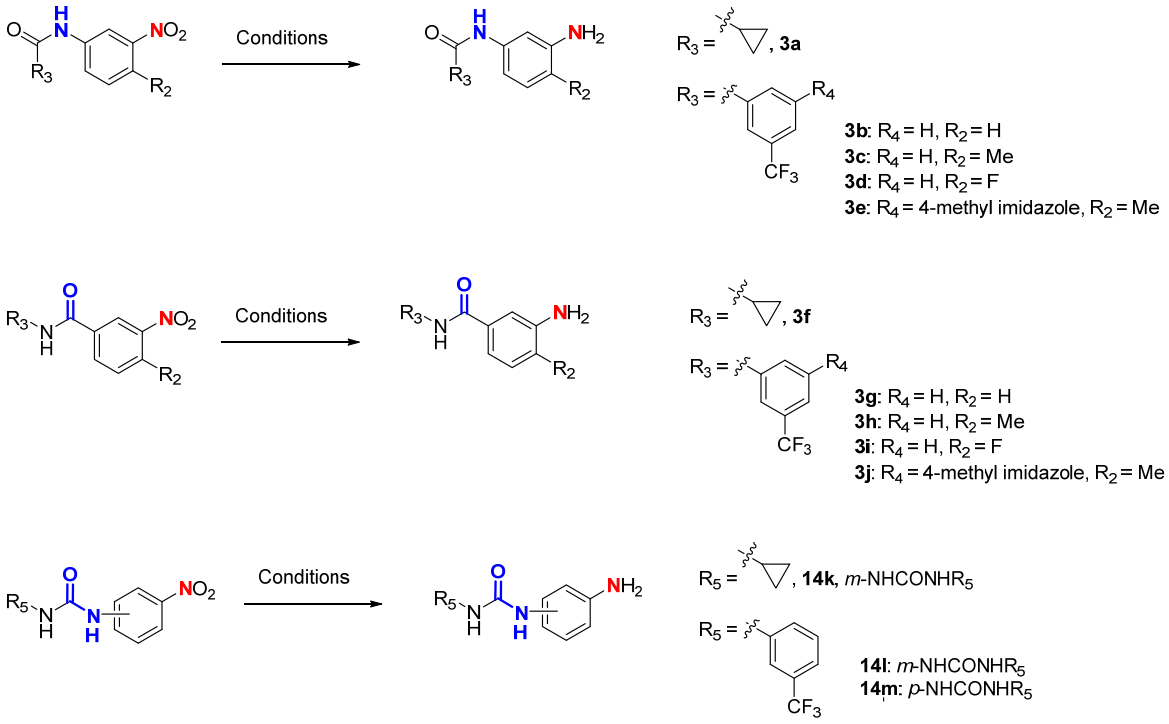

Yellow solid; Yield: 92 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.33 (s, 1H), 9.20 (s, 1H), 8.55 (t, $J$ = 2.1 Hz, 1H), 8.01 (s, 1H), 7.85 (dd, $J$ = 8.1, 2.1 Hz, 1H), 7.75 (dd, $J$ = 8.1, 2.1 Hz, 1H), 7.64-7.62 (m, 1H), 7.58 (t, $J$ = 8.1 Hz, 1H), 7.54 (t, $J$ = 7.8 Hz, 1H), 7.35 (d, $J$ = 7.8 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 152.4, 148.0, 140.6, 140.1, 130.0, 129.9, 129.4 (q, $J$ = 31.5 Hz), 124.5, 124.1 (q, $J$ = 272.3 Hz), 122.2, 118.4 (q, $J$ = 3.7 Hz), 116.5, 114.4 (q, $J$ = 4.0 Hz), 112.3; IR (film): ῡ = 3373, 3082, 1692, 1533, 1337, 1117, 733 cm$^{-1}$; MS (ESI): $m/z$: calcd for $C_{14}H_{10}F_3N_3NaO_3^+$: 348.2, found: 348.2.

**1-(4-Nitrophenyl)-3-(3-(trifluoromethyl)phenyl)urea (14m)[119]**




Yellow solid; Yield: 60 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.16 (s, 2H), 8.01 (s, 2H), 7.60 (d, $J$ = 8.4 Hz, 2H), 7.52 (t, $J$ = 7.9 Hz, 2H), 7.33 (d, $J$ = 7.6 Hz, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 152.5, 140.3, 130.0, 129.5 (q, $J$ = 31.4 Hz), 124.2 (q, $J$ = 272.4 Hz), 122.2, 118.4 (q, $J$ = 3.8 Hz), 114.4 (q, $J$ = 4.1 Hz), 4 C are missing due to overlapping; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.28; IR (neat): ῡ =3356, 3315, 1719, 1658, 1617, 1556, 1494, 1484, 1446, 1325, 1305, 1234, 1175, 1163, 1109, 1070, 931, 884, 849, 796, 751, 697, 653, 495 cm$^{-1}$. MS (ESI), $m/z$: calcd for $C_{14}H_9F_3N_3O_3^-$,[M-H]$^{-1}$: 324.1, found: 324.0.

### III. Reduction of Nitrobenzene Intermediates



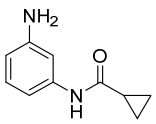

**General procedure A for the reduction of nitro benzene intermediates**

To a solution of nitrophenyl (1 eq) in EtOH (0.1 M) $SnCl_2$ (5 eq) was added. The reaction mixture was heated to 100 °C for 2 h, cooled, and concentrated under reduced pressure. A solution of EtOAc/NaOH (1M) was added to the residue, and the resulting precipitate was filtered off. The organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford the desired products in pure form. This method was used to obtain intermediates **3a**, **3b**, **3f**, **3g**, **3k** and **3l**.

**General procedure B for the reduction of nitro benzene intermediates**

To a solution of nitrophenyl (1 eq) in MeOH (0.3 M) 10% Pd/C (10% wt) was added. The reaction mixture was stirred at 25 °C under a hydrogen balloon for 4-12 hours. The reaction mixture was filtered through a pad of celite and concentrated under reduced pressure obtaining the corresponding anilines in pure form. This method was used to obtain intermediates **3c**[114]**-e, 3h**[118]**-j** and **3m**.[119]

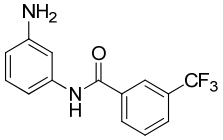
***N*-(3-Aminophenyl)cyclopropanecarboxamide (3a)**



Yellow solid; Yield: 72 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.82 (s, 1H), 6.90-6.87 (m, 2H), 6.68 (d, $J$ = 7.9 Hz, 1H), 6.22 (dd, $J$ = 7.9, 1.3 Hz, 1H), 4.99 (s, 2H), 1.77-1.72 (m, 1H), 076-0.72 (m, 4H); $^{13}$C NMR (125 MHz,

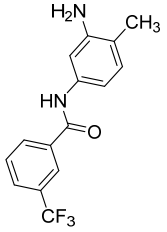
DMSO-$d_6$): δ = 171.1, 148.8, 139.9, 128.7, 108.9, 106.9, 104.7, 14.3, 6.8; IR (film): ῡ = 3440, 3350, 3248, 3013, 1651, 1593, 1456, 1237, 952, 781 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{10}H_{13}N_2O^+$: 177.2, found: 177.0.

**_N_-(3-Aminophenyl)-3-(trifluoromethyl)benzamide (3b)**



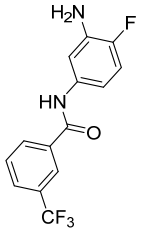
Brown solid; Yield: 81 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.17 (s, 1H), 8.25-8.22 (m, 2H), 7.94 (d, $J$ = 7.8 Hz, 1H), 7.76 (t, $J$ = 7.8 Hz, 1H), 7.09 (t, $J$ = 1.8 Hz, 1H), 6.98 (t, $J$ = 7.9 Hz, 1H), 6.87 (d, $J$ = 7.9 Hz, 1H), 6.35-6.33 (m, 1H), 5.10 (s, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.6, 148.9, 139.3, 136.0, 131.7, 129.5, 129.0 (q, $J$ = 32.1 Hz), 128.7, 127.8 (q, $J$ = 3.6 Hz), 124.1 (q, $J$ = 3.8 Hz), 123.9 (q, $J$ = 272.5 Hz), 110.0, 108.3, 106.1; IR (film): ῡ = 3399, 3246, 1603, 1322, 1263, 1128, 696 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{14}H_{11}F_3N_2NaO^+$: 303.2, found: 303.0.

**_N_-(3-Amino-4-methylphenyl)-3-(trifluoromethyl)benzamide (3c)[114]**



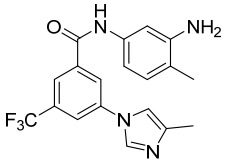
Pale yellow solid; Yield: 77 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.14 (s, 1H), 8.25 (s, 1H), 8.23 (d, $J$ = 7.9 Hz, 1H), 7.93 (d, $J$ = 7.8 Hz, 1H), 7.75 (t, $J$ = 7.8 Hz, 1H), 7.11 (d, $J$ = 1.5 Hz, 1H), 6.84 (dt, $J$ = 8.1, 4.9 Hz, 2H), 4.87 (s, 2H), 2.03 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.7, 146.7, 137.3, 136.2, 131.8, 129.9, 129.8, 129.2 (q, $J$ = 32.0 Hz), 128.0 (q, $J$ = 3.7 Hz), 124.3 (q, $J$ = 3.9 Hz), 124.1 (q, $J$ = 272.5 Hz), 117.4, 109.1, 106.7, 17.1; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.04; IR (neat): ῡ = 3490, 3385, 3271, 1647, 1624, 1615, 1602, 1544, 1518, 1414, 1381, 1326, 1268, 1163, 1111, 1071, 998, 923, 857, 812, 723, 696, 684, 652, 616, 461, 414 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{15}H_{12}F_3N_2O^-$,[M-H]$^{-1}$: 293.1, found: 293.0.

**_N_-(3-Amino-4-fluorophenyl)-3-(trifluoromethyl)benzamide (3d)**



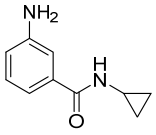
Pale brown solid; Yield: 99 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.24 (s, 1H), 8.25 (s, 1H), 8.23 (d, $J$ = 8.3 Hz, 1H), 7.94 (d, $J$ = 7.7 Hz, 1H), 7.76 (t, $J$ = 7.8 Hz, 1H), 7.29 (dd, $J$ = 8.4, 2.5 Hz, 1H), 6.96 (dd, $J$ = 11.1, 8.7 Hz, 1H), 6.88-6.84 (m, 1H), 5.20 (s, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.9, 147.5 (d, $J$ = 234.3 Hz), 136.4 (d, $J$ = 13.7 Hz), 136.1, 135.3 (d, $J$ = 1.9 Hz), 131.9, 129.8, 129.3 (q, $J$ = 32.0 Hz), 128.1 (q, $J$ = 3.5 Hz), 124.3 (q, $J$ = 3.8 Hz), 124.1 (q, $J$ = 272.5 Hz), 114.7 (d, $J$ = 19.3 Hz), 108.9 (d, $J$ = 4.0 Hz), 108.6 (d, $J$ = 6.4 Hz); $^{19}$F NMR (282 MHz, DMSO): δ = -61.05, -139.62 (ddd, $J$ = 11.3, 8.5, 4.2 Hz); IR (neat): ῡ = 3427, 3300, 1666, 1624, 1555, 1516, 1416, 1328, 1255, 1168, 1107, 1068, 924, 847, 801, 772, 741, 690, 575, 450, 419 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_9F_4N_2O^-$,[M-H]$^{-1}$: 297.0657, found: 297.0654.

**_N_-(3-Amino-4-methylphenyl)-3-(4-methyl-1_H_-imidazol-1-yl)-5(trifluoromethyl)benzamide (3e)**



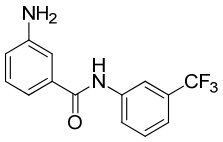
Yellow solid; Yield: 44 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.17 (s, 1H), 8.41 (s, 1H), 8.39 (d, $J$ = 1.4 Hz, 1H), 8.20 (s, 1H), 8.13 (s, 1H), 7.70 (t, $J$ = 1.2 Hz, 1H), 7.09 (d, $J$ = 2.0 Hz, 1H), 6.90 (d, $J$ = 8.44, 1H), 6.83 (dd, $J$ = 8.0, 2.0 Hz, 1H), 4.90 (s, 2H), 2.19 (d, $J$ = 0.8 Hz, 3H), 2.04 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 162.7, 146.6, 138.9, 137.9, 137.7, 137.0, 135.2, 130.7 (q, $J$ = 32.6 Hz), 129.7, 123.5 (q, $J$ = 272.9 Hz), 122.5, 121.8 (q, $J$ = 3.7 Hz), 118.8 (q, $J$ = 3.3 Hz), 117.4, 114.3, 108.9, 106.5, 17.0, 13.5; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.05; IR (neat): ῡ = 3328, 3119, 2969, 1656, 1598, 1543, 1499, 1454, 1426, 1398, 1377, 1330, 1316, 1284, 1253, 1164, 1115, 1078, 1003, 881, 867, 796, 724, 691, 451, 430, 419, 402 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{18}F_3N_4O^+$: 375.1427, found: 375.1430.

**3-Amino-_N_-cyclopropylbenzamide (3f)**



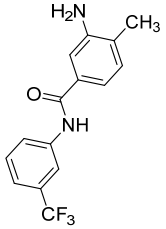
Yellow solid; Yield: 77%; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 8.17 (s, 1H), 7.04 (t, $J$ = 7.7 Hz, 1H), 6.99 (s, 1H), 6.91 (d, $J$ = 7.7 Hz, 1H), 6.68-6.66 (m, 1H), 5.24 (s, 2H), 2.83-2.78 (m, 1H), 0.67-0.63 (m, 2H), 0.55-0.52 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 168.2, 148.3, 135.3, 128.4, 116.3, 114.3, 112.8, 22.9, 5.6; IR (film): ῡ = 3425, 3330, 1628, 1578, 1521, 1485, 1302, 1258, 1022, 746 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{10}H_{12}N_2NaO^+$: 199.2, found: 198.9.

**3-Amino-_N_-(3-(trifluoromethyl)phenyl)benzamide (3g)**



Brown solid; Yield: 77 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.37 (s, 1H), 8.24 (s, 1H), 8.03 (d, $J$ = 8.1 Hz, 1H), 7.57 (t, $J$ = 8.1 Hz, 1H), 7.42 (d, $J$ = 7.8 Hz, 1H), 7.16 (t, $J$ = 7.8 Hz, 1H), 7.11-7.07 (m, 2H), 6.77 (ddd, $J$ = 8.1, 2.2, 0.8 Hz, 1H), 5.33 (s, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.7, 148.7, 140.1, 135.3, 129.7, 129.2 (q, $J$ = 31.5 Hz), 128.7, 124.1 (q, $J$ = 272.3 Hz), 123.5, 119.6 (q, $J$ = 3.8 Hz), 117.0, 116.1 (q, $J$ = 4.1 Hz), 114.6, 112.8; IR (film): ῡ = 3457, 1655, 1539, 1439, 1329, 1120, 793 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{14}H_{11}F_3N_2NaO^+$: 303.2, found: 303.0.

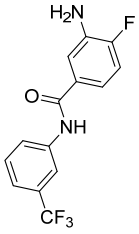
**3-Amino-4-methyl-_N_-(3-(trifluoromethyl)phenyl)benzamide (3h)[118]**



Brown solid; Yield: 99 %; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 10.33 (s, 1H), 8.26 (s, 1H), 8.06 (d, $J$ = 8.3 Hz, 1H), 7.56 (t, $J$ = 8.0 Hz, 1H), 7.41 (d, $J$ = 7.8 Hz, 1H), 7.21 (d, $J$ = 1.6 Hz, 1H), 7.10 (dt, $J$ = 19.4, 4.7 Hz, 2H), 5.10 (s, 2H), 2.13 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.8, 146.8, 140.4, 133.0, 129.9, 129.9, 129.5 (q, $J$ = 31.5 Hz), 125.5, 124.3 (q, $J$ = 272.2 Hz), 123.8, 119.7 (q, $J$ = 3.6 Hz), 116.3 (d, $J$ = 4.0 Hz), 115.4, 113.2, 17.6; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.24; IR (neat): ῡ = 3369, 3320, 1665, 1644, 1576, 1551, 1441,

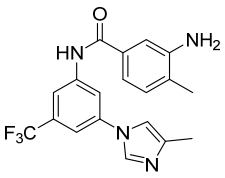
1331, 1318, 1259, 1224, 1175, 1162, 1119, 1105, 1066, 1000, 919, 895, 796, 743, 697, 662, 449, 433 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{15}H_{12}F_3N_2O^-$, $[M-H]^{-1}$: 293.1, found: 293.0.

**3-Amino-4-fluoro-*N*-(3-(trifluoromethyl)phenyl) benzamide (3i)**



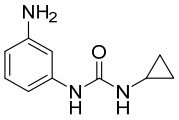
Pale brown solid; Yield: 92 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.39 (s, 1H), 8.22 (s, 1H), 8.02 (d, *J* = 8.4 Hz, 1H), 7.58 (t, *J* = 7.9 Hz, 1H), 7.43 (d, *J* = 7.8 Hz, 1H), 7.36 (d, *J* = 7.6 Hz, 1H), 7.17-7.11 (m, 2H), 5.41 (s, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 165.8, 152.6 (d, *J* = 243.0 Hz), 140.1, 136.6 (d, *J* = 13.4 Hz), 131.2 (d, *J* = 2.6 Hz), 129.8, 129.4 (q, *J* = 31.7 Hz), 124.2 (q, *J* = 272.3 Hz), 123.7, 119.8 (q, *J* = 3.3 Hz), 116.3 (q, *J* = 4.0 Hz), 115.9 (d, *J* = 5.9 Hz), 115.6 (d, *J* = 7.4 Hz), 114.7 (d, *J* = 19.2 Hz); $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.25, -130.26 (dd, *J* = 16.0, 8.7 Hz); IR (neat): $\tilde{\upsilon}$ = 3406, 3253, 3206, 3088, 1661, 1604, 1561, 1511, 1444, 1431, 1336, 1324, 1282, 1261, 1198, 1179, 1162, 1101, 1066, 919, 899, 870, 849, 791, 771, 752, 724, 694, 660, 642, 556, 450 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_9F_4N_2O^-$, $[M-H]^{-1}$: 297.0657, found: 297.0656.

**3-Amino-4-methyl-*N*-(3-(4-methyl-1*H*-imidazol-1-yl)-5(trifluoromethyl))benzamide (3j)**



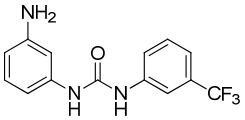
Yellow solid; Yield: 52 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.72 (s, 1H), 9.41 (s, 1H), 8.55 (s, 1H), 8.27 (s, 1H), 7.94 (s, 1H), 7.86 (s, 1H), 7.31 (s, 1H), 7.27 (d, *J* = 8.0 Hz, 1H), 7.14 (d, *J* = 7.8 Hz, 1H), 2.34 (s, 3H), 2.17 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 145.3, 141.8, 136.5, 134.4, 133.3, 132.3, 130.8 (q, *J* = 32.6 Hz), 130.0, 126.9, 123.5 (q, *J* = 272.7 Hz), 116.5, 116.5, 116.3, 116.1, 114.2, 112.8, 17.6, 10.9; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.42; IR (neat): $\tilde{\upsilon}$ = 3328, 3229, 2795, 1667, 1613, 1541, 1508, 1480, 1441, 1398, 1381, 1304, 1231, 1201, 1171, 1123, 1088, 998, 929, 876, 850, 823, 745, 707, 693, 634, 539, 424 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{18}F_3N_4O^+$: 375.1427, found: 375.1432.

**1-(3-Aminophenyl)-3-cyclopropylurea (3k)**



Yellow solid; Yield: 53 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 7.92 (s, 1H), 6.82 (t, *J* = 7.9 Hz, 1H), 6.68 (t, *J* = 1.9 Hz, 1H), 6.52-6.50 (m, 1H), 6.22 (d, *J* = 1.8 Hz, 1H), 6.13-6.11 (m, 1H), 4.91 (s, 2H), 2.52-2.48 (m, 1H), 0.63-0.59 (m, 2H), 0.38-0.35 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 165.2, 147.6, 135.7, 133.5, 129.9, 125.7, 121.7, 23.1, 5.6; IR (film): $\tilde{\upsilon}$ = 3332, 1607, 1548, 1495, 1218, 1022, 760 $cm^{-1}$; MS (ESI): *m/z*: calcd for $C_{10}H_{13}N_3NaO^+$: 214.2, found: 214.0.

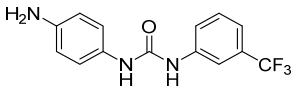
**1-(3-Aminophenyl)-3-(3-(trifluoromethyl)phenyl)urea (3l)**



White solid; Yield: 92 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.89 (s, 1H), 8.45 (s, 1H), 8.03 (s, 1H), 7.53-7.48 (m, 2H), 7.28 (d, *J* = 6.9 Hz, 1H), 6.90 (t, *J* = 7.9 Hz, 1H), 6.81 (t, *J* = 2.0 Hz, 1H), 6.54 (dd, *J* =

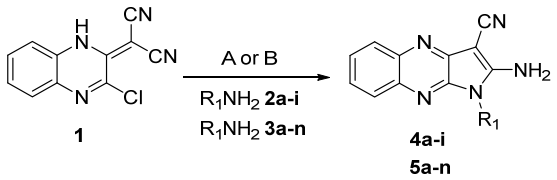
7.9, 2.0 Hz, 1H), 6.21 (dd, $J$ = 7.9, 2.0 Hz, 1H), 5.03 (s, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 152.2, 149.1, 140.6, 139.8, 129.8, 129.4 (q, $J$ = 31.3 Hz), 128.9, 124.1 (q, $J$ = 272.3 Hz), 121.5, 117.7 (q, $J$ = 3.8 Hz), 113.8 (q, $J$ = 4.1 Hz), 108.3, 106.2, 103.9; IR (film): ῡ = 3306, 1633, 1557, 1333, 1110, 769 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{14}H_{12}F_3N_3NaO^+$: 318.3, found: 318.1.

**1-(4-Aminophenyl)-3-(3-(trifluoromethyl)phenyl)urea (3m)[119]**



Brown solid; Yield: 69 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 9.41 (s, 2H), 8.00 (s, 2H), 7.61-7.57 (m, 3H), 7.54-7.50 (m, 2H), 7.33-7.28 (m, 3H); $^{13}$C NMR (125 MHz, DMSO): δ = 152.7, 152.6, 140.8, 140.3, 133.1, 129.9, 129.9, 129.6 (q, $J$ = 24.9 Hz), 124.2 (q, $J$ = 272.3 Hz), 122.1, 121.6, 120.3, 118.4 (q, $J$ = 3.8 Hz), 114.3 (q, $J$ = 4.0 Hz); $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.30; IR (neat): ῡ = 3304, 2873, 1599, 1656, 1615, 1556, 1515, 1493, 1448, 1424, 1329, 1311, 1230, 1215, 1164, 1119, 1096, 1068, 884, 794, 741, 699, 655, 512 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{13}F_3N_3O^+$: 296.26, found: 296.0.

## IV. Cyclization of 2-(3-chloroquinoxalin-2-yl)malononitrile




1 4a-i 5a-n

**General procedure A for the cyclization of 2-(3-chloroquinoxalin-2-yl)malononitrile**

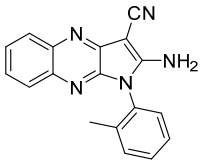
To a solution of 2-(3-chloroquinoxalin-2-yl)malononitrile (1 eq) in 1 : 1 EtOH – toluene solution (0.09 M) the primary amine (4 eq for **4a-i**, **5a**, **5f**, **5k** and **5l**, 2 eq for **5b-d**, **5g-i** and **5m** and 1 eq for **5n**) was added. The reaction mixture was heated to 130-160 °C for 4-12 h and cooled down. The formed solid was filtered off and washed with $Et_2O$ and EtOH to obtain the desired compounds **4a**, **4d-i**, **5a**, **5d**, **5i**, **5f**, **5k** and **5l** in pure form. Compounds **5c** and **5i** were purified by preparative thin layer chromatography and **4b**, **4c** and **5n** by flash column chromatography (hex/EtOAc, 1:1).

**General procedure B for the cyclization of 2-(3-chloroquinoxalin-2-yl)malononitrile**

To a solution of 2-(3-chloroquinoxalin-2-yl)malononitrile (1.0 eq) in DMF (0.10 M) the primary amine (1.0 eq for **5e**, **5j** and 1.2 eq for **5m**) was added. The reaction mixture was heated to 80 °C for 12 h, concentrated and purified by flash column chromatography

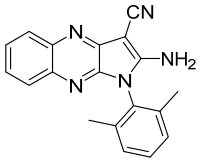
(hexane/EtOAc 1:1 to EtOAc/MeOH/$Et_3N$, 90:9:1 for **5m** and EtOAc/MeOH, 9:1 for **5e**, **5j**).

**2-Amino-1-(*o*-tolyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4a)**



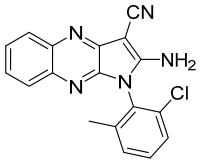
Yellow solid; Yield: 73 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 8.27 (s, 2H), 7.94 (d, $J$ = 8.3 Hz, 1H), 7.74 (d, $J$ = 8.1 Hz, 1H), 7.60-7.44 (m, 6H), 2.02 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.7, 143.6, 142.4, 140.3, 137.3, 137.3, 131.4, 130.9, 130.2, 129.9, 127.5, 127.5, 127.1, 126.8, 125.6, 115.2, 60.3, 17.1; IR (film): $\tilde{\upsilon}$ = 3356, 3299, 3065, 2205, 1638, 1551, 1482, 1432, 1398, 753, 730, 597 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{18}H_{14}N_5^+$: 300.1244, found: 300.1241.

**2-Amino-1-(2,6-dimethylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4b)**



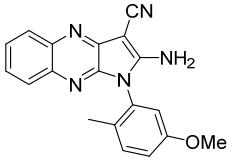
Orange solid; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.34 (s, 2H), 7.95 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.76 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.59 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.48 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.45-7.42 (m, 1H), 7.33 (d, $J$ = 7.6 Hz, 2H), 1.95 (s, 6H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.2, 143.5, 141.5, 140.4, 137.5, 137.4, 130.0, 129.8, 128.8, 127.6, 127.1, 126.9, 125.7, 115.0, 60.2, 17.3; IR (film): $\tilde{\upsilon}$ = 3306, 3059, 2986, 2204, 1665, 1606, 1566, 763, 754, 595 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{19}H_{16}N_5^+$: 314.1400, found: 314.1400.

**2-Amino-1-(2-chloro-6-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4c)**



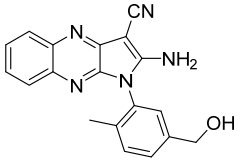
Orange solid; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.47 (s, 2H), 7.94 (d, $J$ = 8.2 Hz, 1H), 7.76 (d, $J$ = 8.3 Hz, 1H), 7.61-7.55 (m, 3H), 7.51-7.47 (m, 2H), 2.08 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.3, 143.6, 141.5, 140.7, 140.5, 137.2, 133.5, 131.5, 130.0, 128.5, 127.9, 127.5, 127.0, 126.9, 125.7, 115.0, 60.4, 17.6; IR (film): $\tilde{\upsilon}$ = 3332, 3154, 2923, 2208, 1656, 1559, 1462, 1209, 1025, 762 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{18}ClH_{13}N_5^+$: 334.0854, found: 334.0853.

**2-Amino-1-(5-methoxy-2-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4d)**



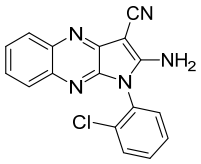
Yellow solid; Yield: 53 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.30 (s, 2H), 7.94 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.77 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.58 (ddd, $J$ = 8.3, 6.9, 1.3 Hz, 1H), 7.48 (ddd, $J$ = 8.3, 6.9 Hz, 1.3 Hz, 1H), 7.41 (d, $J$ = 8.5 Hz, 1H), 7.13-7.09 (m, 2H), 3.77 (s, 3H), 1.91 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.6, 158.3, 143.6, 142.2, 140.3, 137.3, 131.9, 131.4, 128.7, 127.5, 127.0, 126.8, 125.6, 116.2, 115.2, 115.1, 60.2, 55.4, 16.2; IR (film): $\tilde{\upsilon}$ = 3315, 3144, 3061, 2208, 1636, 1546, 1506, 1447, 1313, 1209, 1030, 820, 760, 735 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{19}H_{16}N_5O^+$: 330.1349, found: 330.1348.

**2-Amino-1-(5-(hydroxymethyl)-2-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4e)**



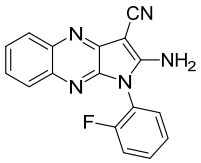

Yellow solid; Yield: 74 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 8.28 (s, 2H), 7.94 (d, $J$ = 8.0 Hz, 1H), 7.75 (d, $J$ = 7.8 Hz, 1H), 7.58 (ddd, $J$ = 8.0, 7.1, 1.0 Hz, 1H), 7.49-7.47 (m, 3H), 7.34 (s, 1H), 5.33 (t, $J$ = 5.5 Hz, 1H), 4.56 (d, $J$ = 5.5 Hz, 2H), 1.98 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.7, 143.6, 142.3, 142.1, 140.3, 137.2, 135.3, 131.1, 130.5, 128.1, 127.5, 127.4, 127.1, 126.8, 125.7, 115.1, 62.0, 60.3, 16.8; IR (film): ῡ = 3343, 3309, 3124, 2208, 1637, 1550, 1446, 766, 599 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{19}H_{16}N_5O^+$: 330.1349, found: 330.1346.

**2-Amino-1-(2-chlorophenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4f)**



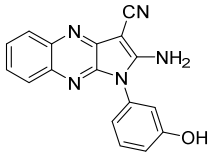

Orange solid; Yield: 37 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 8.45 (s, 2H), 7.95 (dd, $J$ = 8.3, 1.0 Hz, 1H), 7.79 (dd, $J$ = 8.1, 1.3 Hz, 1H), 7.77-7.74 (m, 2H), 7.68 (td, $J$ = 7.6, 1.7 Hz, 1H), 7.64-7.58 (m, 2H), 7.49 (ddd, $J$ = 8.3, 7.1, 1.3 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.6, 143.5, 142.3, 140.4, 137.1, 133.1, 132.2, 132.0, 130.5, 129.5, 128.8, 127.5, 127.0, 126.9, 125.7, 114.9, 60.4; IR (film): ῡ = 3347, 3301, 2210, 1641, 1551, 1477, 1432, 1399, 743 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{17}H_{10}ClN_5Na^+$: 342.7, found: 342.1.

**2-Amino-1-(2-fluorophenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4g)**



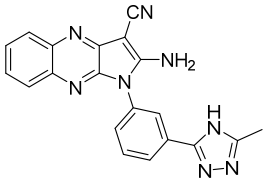

Yellow solid; Yield: 79 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 8.43 (s, 2H), 7.95 (dd, $J$ = 8.3, 1.1 Hz, 1H), 7.76 (dd, $J$ = 8.3, 1.2 Hz, 1H), 7.72-7-67 (m, 2H), 7.62-7.54 (m, 2H), 7.51-7.46 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.7, 158.4 (d, $J$ = 250.4 Hz), 143.5, 142.4, 140.4, 137.1, 132.3 (d, $J$ = 8.0 Hz), 131.7, 127.5, 127.1, 127.0, 125.8, 125.7 (d, $J$ = 3.6 Hz), 119.3 (d, $J$ = 13.2 Hz), 117.1 (d, $J$ = 19.1 Hz), 114.8, 60.7; IR (film): ῡ = 3308, 3109, 2206, 1639, 1555, 1436, 1213, 760 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{17}H_{10}FN_5Na^+$: 326.3, found: 326.2.

**2-Amino-1-(3-hydroxyphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4h)**



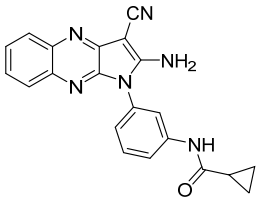

Yellow solid; Yield: 73 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 9.93 (s, 1H), 8.26 (s, 2H), 7.93 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.77 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.58 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.49 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.42 (t, $J$ = 8.0 Hz, 1H), 6.99 (ddd, $J$ = 8.0, 2.1, 0.8 Hz, 1H), 6.95 (ddd, $J$ = 8.0, 2.1, 0.8 Hz, 1H), 6.90 (t, $J$ = 2.1 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.6, 158.3, 143.5, 142.6, 140.3, 137.2, 132.6, 130.5, 127.5, 127.0, 126.7, 125.6, 119.1, 116.6, 115.6, 115.1, 60.4; IR (film): ῡ = 3437, 3349, 2952, 1654, 1554, 1433, 1282, 824, 767 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_{17}H_{11}N_5NaO^+$: 324.3, found: 324.2.

**2-Amino-1-(3-(5-methyl-4H-1,2,4-triazol-3-yl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (4i)**



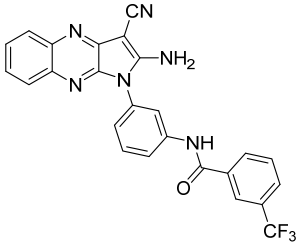
Orange solid; Yield: 44 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 13.77 (s, 1H), 8.37 (s, 2H), 8.19 (d, $J$ = 7.9 Hz, 1H), 8.08 (s, 1H), 7.94 (d, $J$ = 8.3 Hz, 1H), 7.76-7.71 (m, 2H), 7.60-7.57 (m, 2H), 7.48 (t, $J$ = 7.5 Hz, 1H), 2.41 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.8, 153.7, 143.6, 142.8, 140.3, 137.1, 133.2, 132.3, 130.3, 129.1, 127.5, 127.0, 126.7, 126.5, 126.0, 125.6, 115.1, 60.7, 11.6, 1C is missing due to overlapping; IR (film): ῦ = 3452, 3312, 3173, 2219, 1659, 1561, 1466, 1410, 1045, 740 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{20}H_{16}N_8^+$: 367.1414, found: 367.1417.

***N*-(3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)phenyl)cyclopropanecarboxamide (5a)**



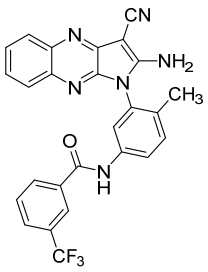
Yellow solid; Yield: 65 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.56 (s, 1H), 8.48 (s, 2H), 7.94 (d, $J$ = 8.1 Hz, 1H), 7.82-7.78 (m, 3H), 7.62-7.49 (m, 3H), 7.21 (d, $J$ = 7.6 Hz, 1H), 1.83 (quint, $J$ = 6.1 Hz, 1H), 0.82-0.81 (m, 4H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 171.9, 159.9, 143.2, 142.7, 140.6, 138.8, 136.9, 131.8, 130.1, 127.6, 127.1, 126.0, 125.8, 123.1, 119.8, 118.8, 114.7, 60.8, 14.5, 7.2; IR (film): ῦ = 3239, 2208, 1653, 1551, 1450, 1201, 954, 756 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{21}H_{17}N_6O^+$: 369.1458, found: 369.1459.

***N*-(3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)phenyl)-3-(trifluoromethyl)benzamide (5b)**



Yellow solid; Yield: 78 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.76 (s, 1H), 8.36-8.28 (m, 4H), 8.05 (d, $J$ = 8.4 Hz, 1H), 8.00-7.94 (m, 3H), 7.83-7-77 (m, 2H), 6.66 (t, $J$ = 8.1 Hz, 1H), 7.60-7.57 (m, 1H), 7.51-7.48 (m, 1H), 7.34 (d, $J$ = 7.9 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.1, 159.6, 143.5, 142.7, 140.3, 140.1, 137.1, 135.4, 132.0, 131.8, 130.1, 129.8, 129.1 (q, $J$ = 32.1 Hz), 128.3 (q, $J$ = 3.5 Hz), 127.5, 127.0, 126.8, 125.6, 124.3, 124.1 (q, $J$ = 3.8 Hz), 123.8 (q, $J$ = 272.5Hz), 121.2, 120.2, 115.0, 60.6; IR (film): ῦ = 3326, 3155, 2210, 1663, 1561, 1426, 1115, 660 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{25}H_{15}F_3N_6NaO^+$: 495.1152, found: 495.1146.

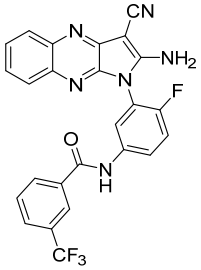
***N*-(3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)-4-methyl phenyl)-3-(trifluoromethyl)benzamide (5c)**



Yellow solid; Yield: 30 %; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 10.67 (s, 1H), 8.36 (br, 2H), 8.31(s, 1H), 8.27 (d, $J$ = 7.9 Hz, 1H), 8.03-7.91 (m, 3H), 7.86 (d, $J$ = 1.8 Hz, 1H), 7.83-7.74 (m, 2H), 7.59 (t, $J$ = 7.1 Hz, 1H), 7.55-7.45 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.0,

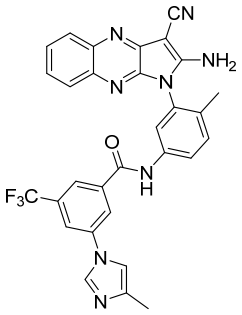
159.6, 143.6, 142.3, 140.4, 138.1, 137.3, 135.5, 132.7, 131.9, 131.6, 130.8, 129.8, 129.2 (q, $J$ = 32.1 Hz), 128.3 (q, $J$ = 3.1 Hz), 127.6, 127.2, 126.9, 125.8, 124.17 (q, $J$ = 3.8 Hz), 123.95 (q, $J$ = 272.5 Hz), 122.1, 121.2, 115.1, 60.43, 16.68; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.07; IR (neat): ῦ = 3326, 3166, 2212, 1656, 1596, 1565, 1508, 1485, 1445, 1391, 1326, 1309, 1256, 1203, 1166, 1124, 1092, 1073, 1037, 817, 755, 693, 650, 599, 477, 411 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{26}H_{18}F_3N_6O^+$: 487.1489, found: 487.1492.

**_N_-(3-(2-Amino-3-cyano-1_H_-pyrrolo [2,3-_b_] quinoxalin-1-yl)-4-fluorophenyl)-3-(trifluoromethyl)benzamide (5d)**



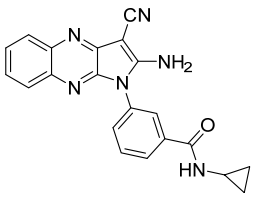
Yellow solid; Yield: 57 %;$^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.87 (s, 1H), 8.77 (br, 2H), 8.33 (s, 1H), 8.31 (d, $J$ = 7.7 Hz, 1H), 8.12 (dd, $J$ = 6.5, 2.7 Hz, 1H), 8.11-8.06 (m, 1H), 8.00-7.96 (m, 2H), 7.82-7.79 (m, 2H), 7.65-7.58 (m, 2H), 7.53 (dd, $J$ = 8.2, 6.9 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.1, 160.0, 155.6, 153.7, 143.0, 142.6, 136.9, 136.2, 135.3, 131.9, 129.8, 129.3, 129.1, 128.4, 128.2, 127.7, 127.5, 126.2, 125.0, 124.2, 122.9, 118.8, 117.3, 117.2, 114.4, 61.3; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.07, -125.90 – -126.02 (m); IR (neat): ῦ = 3338, 3076, 2230, 1649, 1624, 1552, 1511, 1471, 1390, 1329, 1253, 1223, 1120, 1073, 917, 834, 809, 755, 693, 658, 596, 510, 419 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{15}F_4N_6O^+$: 491.1238, found: 491.1240.

**_N_-(3-(2-Amino-3-cyano-1_H_-pyrrolo[2,3-_b_]quinoxalin-1-yl)-4-methylphenyl)-3-((4-methyl-1_H_-imidazol-1-yl)-5-(trifluoromethyl)benzamide) (5e)**



Brown solid; Yield: 24 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.70 (s, 1H), 8.44-8.38 (m, 5H), 8.25 (s, 1H), 8.18 (s, 1H), 7.97 (dd, $J$ = 13.0, 9.2 Hz, 1H), 7.83 (s, 1H), 7.77 (d, $J$ = 8.2 Hz, 1H), 7.69 (s, 1H), 7.60 (t, $J$ = 7.1 Hz, 1H), 7.55 (d, $J$ = 8.5 Hz, 1H), 7.49 (t, $J$ = 7.5 Hz, 1H), 2.18 (s, 3H), 2.01 (s, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ = 163.2, 159.6, 143.6, 142.3, 140.4, 139.0, 137.8, 137.3, 137.3, 135.3, 133.0, 131.7, 130.9, 127.6, 127.2, 126.9, 125.8, 122.6, 122.2, 121.8, 121.4, 115.1, 114.2, 60.4, 16.7, 13.6; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.07; IR (neat): ῦ = 3299, 3066, 2925, 2204, 1651, 1603, 1565, 1538, 1498, 1463, 1393, 1378, 1317, 1281, 1251, 1176, 1125, 1080, 1011, 880, 817, 755, 691, 664, 598, 446 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{30}H_{22}F_3N_8O^+$: 567.1863, found: 567.1862.

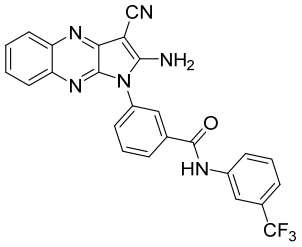
**3-(2-Amino-3-cyano-1_H_-pyrrolo[2,3-_b_]quinoxalin-1-yl)-_N_-cyclopropylbenzamide (5f)**



Yellow solid; Yield: 55 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.52 (d, $J$ = 4.0 Hz, 1H), 8.36 (s, 2H), 8.07-8.05 (m, 1H), 7.99 (s, 1H), 7.94 (dd, $J$ = 8.3, 1.1 Hz, 1H), 7.76-7.71 (m, 3H), 7.59 (ddd, $J$ = 8.3, 7.0, 1.1 Hz, 1H), 7.49 (ddd, $J$ = 8.3, 7.0, 1.1 Hz, 1H), 2.91-2.86 (m,

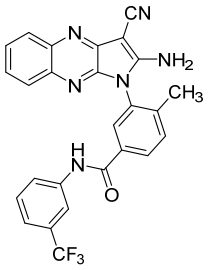
1H), 0.73-0.69 (m, 2H), 0.59-0.56 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.1, 159.7, 143.5, 142.7, 140.3, 137.1, 135.8, 131.9, 131.7, 129.9, 128.3, 127.5, 127.0, 126.8, 125.7, 115.0, 60.6, 23.0, 5.6, 1C is missing due to overlapping; IR (film): ῡ = 3285, 3116, 2210, 1644, 1563, 1444, 1295, 1143, 747 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{21}H_{17}N_6O^+$: 369.1458, found: 369.1456.

**3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)-*N*-(3-(trifluoromethyl)phenyl)benzamide (5g)**



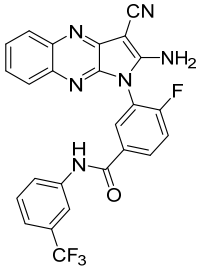
Yellow solid; Yield: 31 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.61 (s, 1H), 8.43 (s, 2H), 8.24-8.20 (m, 3H), 8.08 (d, *J* = 8.3 Hz, 1H), 7.96 (dd, *J* = 8.3, 1.0 Hz, 1H), 7.87-7.82 (m, 2H), 7.77 (dd, *J* = 8.3, 1.0 Hz, 1H), 7.64-7-58 (m, 2H), 7.51-7.47 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.6, 159.7, 143.5, 142.7, 140.4, 139.6, 137.1, 135.8, 132.6, 132.1, 130.1, 129.9, 129.3 (q, *J* = 31.6 Hz), 128.9, 128.0, 127.5, 127.1, 126.9, 125.7, 124.0 (q, *J* = 272.3 Hz), 123.7, 120.1 (q, *J* = 3.8 Hz), 116.3 (q, *J* = 4.0 Hz), 115.0, 60.7; IR (film): ῡ = 3346, 3184, 2211, 1660, 1561, 1443, 1331, 1121, 596 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{25}H_{16}F_3N_6O^+$: 473.1332, found: 473.1334.

**3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)-4-methyl-*N*-(3-(trifluoromethyl)phenyl)benzamide (5h)**



Yellow solid; Yield: 33 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.53 (s, 1H), 8.43 (br, 2H), 8.21 (s, 1H), 8.19 (dd, *J* = 8.2, 1.6 Hz, 1H), 8.13 (s, 1H), 8.08 (d, *J* = 8.4 Hz, 1H), 7.96 (d, *J* = 7.6 Hz, 1H), 7.77 (d, *J* = 7.6 Hz, 1H), 7.72 (d, *J* = 8.1 Hz, 1H), 7.62-7.59 (m, 2H), 7.51-7.45 (m, 2H), 2.12 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 164.5, 159.9, 143.8, 142.5, 142.2, 140.6, 139.9, 137.4, 133.6, 131.9, 131.3, 130.1, 129.9, 129.8, 129.6, 129.5, 129.2, 127.8, 127.4, 127.2, 126.0, 125.4, 124.0, 123.2, 120.3, 116.7, 115.3, 60.8, 17.5; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.27; IR (neat): ῡ = 3323, 3130, 2922, 2360, 2205, 1651, 1555, 1443, 1332, 1260, 1165, 1121, 1069, 796, 759, 696, 667, 659, 598, 418 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{26}H_{18}F_3N_6O^+$: 487.1489, found: 487.1493.

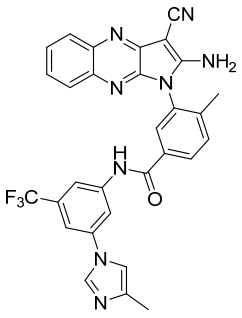
**3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)-4-fluoro-*N*-(3-(trifluoromethyl)phenyl)benzamide (5i)**



Yellow solid; Yield: 65 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.71 (s, 1H), 8.74 (br, 2H), 8.39-8.35 (m, 2H), 8.23 (s, 1H), 8.09 (d, *J* = 9.0 Hz, 1H), 7.97 (d, *J* = 8.4 Hz, 1H), 7.81-7.76 (m, 2H), 7.65-7.60 (m, 2H), 7.54-7.47 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 163.8, 161.8, 160.0, 159.6, 143.1, 139.7, 137.0, 132.4, 131.9, 131.8, 130.0, 129.5, 129.3, 127.7, 127.5, 126.2, 125.2, 123.9, 123.0, 120.3, 117.6, 117.4, 116.5, 116.5, 114.7,

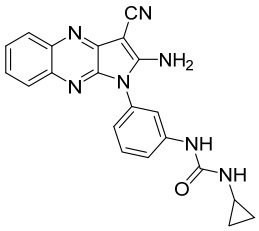
61.3; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.27, -115.32 – -115.52 (m); IR (neat): ῡ = 3258, 3121, 2226, 1884, 1686, 1652, 1618, 1581, 1563, 1515, 1493, 1471, 1389, 1330, 1277, 1212, 1172, 1120, 1095, 1073, 887, 844, 796, 761, 748, 694, 594, 517, 510, 479, 443 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{15}F_4N_6O^+$: 491.1238, found: 491.1244.

**3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)-4-methyl-*N*-(4-methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)benzamide) (5j)**



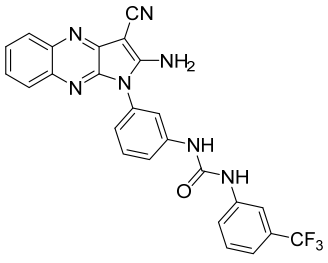
Orange solid; Yield: 21 %; $^1$H-NMR (400 MHz, DMSO-$d_6$): δ = 10.65 (s, 1H), 8.43 (s, 2H), 8.27 (s, 1H), 8.19-8.21 (m, 2H), 8.14 (s, 2H), 7.97 (d, *J* = 8.5 Hz, 1H), 7.78-7.73 (m, 3H), 7.60 (t, *J* = 7.1 Hz, 1H), 7.51-7.47 (m, 2H), 2.17 (s, 3H), 2.13 (s, 3H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ = 164.5, 159.7, 143.6, 142.3, 142.3, 141.1, 140.5, 138.9, 137.9, 137.3, 134.9, 133.0, 131.8, 131.2, 131.0, 130.7, 130.5, 129.7, 129.3, 127.6, 127.2, 127.0, 125.9, 124.7, 115.0, 115.0, 114.2, 60.6, 17.4, 13.5; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.34; IR (neat): ῡ = 3303, 2924, 2855, 2202, 1683, 1653, 1614, 1565, 1499, 1479, 1444, 1407, 1325, 1309, 1269, 1241, 1227, 1193, 1170, 1123, 1103, 1082, 1021, 1008, 930, 863, 808, 758, 750, 694, 671, 627, 597, 523, 473, 446, 413, 402 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{30}H_{22}F_3N_8O^+$: 567.1863, found: 567.1861.

**1-(3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)phenyl)-3-cyclopropylurea (5k)**



Yellow solid; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.60 (s, 1H), 8.28 (s, 2H), 7.93 (d, *J* = 8.3 Hz, 1H), 7.77 (d, *J* = 8.2 Hz, 1H), 7.61-7.57 (m, 3H), 7.50-7.46 (m, 2H), 7.07 (d, *J* = 7.9 Hz, 1H), 6.50 (s, 1H), 2.56-2.52 (m, 1H), 0.65-0.61 (m, 2H), 0.42-0.39 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.6, 155.8, 143.5, 142.6, 141.7, 140.3, 137.2, 132.0, 129.9, 127.5, 127.0, 126.7, 125.6, 121.2, 118.7, 117.7, 115.1, 60.4, 22.3, 6.3; IR (film): ῡ = 3321, 3160, 2923, 2207, 1695, 1537, 1449, 1244, 1031, 754 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{21}H_{18}N_7O^+$: 384.1567, found: 384.1567.

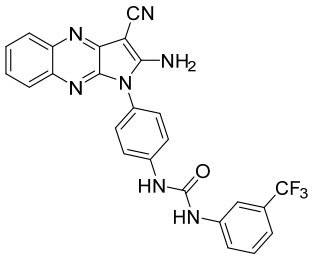
**1-(3-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)phenyl)-3-(3-(trifluoromethyl)phenyl)urea (5l)**



Yellow solid; Yield: 54 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.19 (s, 1H), 9.14 (s, 1H), 8.33 (s, 2H), 8.02 (s, 1H), 7.94 (d, *J* = 8.1 Hz, 1H), 7.77 (d, *J* = 8.1 Hz, 1H), 7.70-7.67 (m, 2H), 7.60-7.47 (m, 5H), 7.32 (d, *J* = 7.5 Hz, 1H), 7.18 (d, *J* = 7.3 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 159.6, 152.4, 143.5, 142.7, 140.7, 140.3, 137.2, 132.2, 130.1, 129.8, 129.4 (q, *J* = 31.4 Hz), 127.5, 127.0, 126.7, 125.6, 124.1 (q, *J* = 272.4 Hz), 122.3, 121.8, 119.3, 118.3, 118.2 (q, *J* = 3.4 Hz), 115.1, 114.1 (q, *J* = 4.4 Hz), 60.5, 1C is missing due to overlapping;

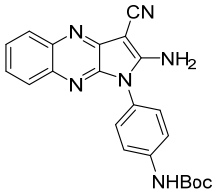
IR (film): ῦ = 3322, 3157, 2208, 1663, 1560, 1446, 1320, 1163, 1117, 754 $cm^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{25}H_{17}F_3N_7O^+$: 488.1441, found: 488.1442.

**1-(4-(2-Amino-3-cyano-1*H*-pyrrolo[2,3-*b*]quinoxalin-1-yl)phenyl)-3-(3-(trifluoromethyl)phenyl)urea (5m)**



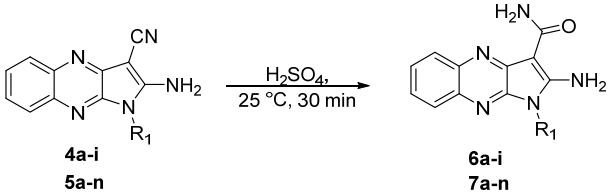
Yellow solid; Yield: 61 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.18 (s, 1H), 9.15 (s, 1H), 8.28 (br, 2H), 7.93 (dd, *J* = 8.4, 1.0 Hz, 1H), 7.76 (dd, *J* = 8.3, 1.2 Hz, 1H), 7.72 (d, *J* = 8.8 Hz, 2H), 7.63 (d, *J* = 7.5 Hz, 1H), 7.60 – 7.51 (m, 3H), 7.47 (d, *J* = 8.8 Hz, 2H), 7.34 (d, *J* = 7.6 Hz, 1H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ = 213.8, 160.0, 152.5, 143.7, 143.0, 140.6, 140.4, 140.4, 137.3, 130.0, 129.7, 129.5, 129.4, 127.6, 127.1, 126.7, 125.6, 125.5, 125.5, 122.8, 122.0, 119.4, 118.3 (q, *J* = 4.5 Hz), 115.2, 114.2 (q, *J* = 3.2 Hz), 60.4; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.27; IR (neat): ῦ = 3282, 3066, 2597, 2225, 1655, 1622, 1602, 1551, 1514, 1493, 1446, 1336, 1313, 1232, 1201, 1174, 1166, 1120, 1069, 834, 793, 757, 696, 668, 659, 596, 510, 483,474, 420 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{17}F_3N_7O^+$: 488.1441, found: 488.1441.

**2-Amino-1-(*N*-Boc-*p*-phenylamine)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carbonitrile (5n)**



Yellow solid; 46 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 9.66 (s, 1H), 8.24 (br, 2H), 7.92 (dd, *J* = 8.5, 1.3 Hz, 1H), 7.74 (dd, *J* = 8.5, 1.3 Hz, 1H), 7.68 (d, *J* = 8.7 Hz, 2H), 7.59-7.55 (m, 1H), 7.48 (ddd, *J* = 8.3, 7.1, 1.6 Hz, 1H), 7.42 (d, *J* = 8.8 Hz, 2H), 1.52 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 160.0, 152.8, 143.7, 143.0, 140.7, 140.4, 137.3, 129.4, 127.6, 127.1, 126.8, 125.7, 125.4, 119.1, 115.2, 79.5, 60.4, 28.1; IR (neat): ῦ = 3349, 3256, 3160, 2214, 1700, 1671, 1592, 1565, 1502, 1482, 1443, 1413, 1394, 1369, 1314, 1238, 1214, 1166, 1058, 1028, 1020, 840, 752, 624, 599, 528, 418 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{22}H_{21}N_6O_2^+$: 401.1721, found: 401.1720.

### V. Synthesis of quinoxaline carboxamides



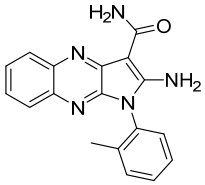

**General procedure for the hydrolysis of carbonitriles**

A solution of the carbonitrile (1.0 eq) in sulfuric acid (0.15 M) was stirred for 30 min at 25 °C. The reaction was poured into ice cold water and basified with a solution of

ammonium hydroxide (25%) to pH 7. The resulting solid was filtered off and washed with cold water obtaining the final products. The crude product of **6e** was then heated in acetone (5.0 mL) at 160 °C for 1 h. After cooling to room temperature, the solid was filtered off and washed with hexane to afford the desired product in pure form.

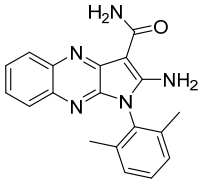
In the case of **6a-d**, **6f-i** and **7a**, **7f**, **7k** the residue was redissolved in EtOAc and extracted with a saturated solution of $NaHCO_3$ affording the corresponding product in pure form. **7l** and **7f** were further recrystallized in acetone and **7k** by column chromatography (EtOAc).

**2-Amino-1-(*o*-tolyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6a)**



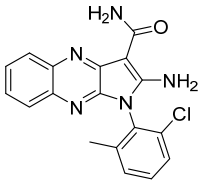
Yellow solid; Yield: 74 %; mp 152-155 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.05 (s, 2H), 7.95 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.77-7.75 (m, 2H), 7.59-7.53 (m, 3H), 7.48-7.44 (m, 3H), 7.39 (d, $J$ = 2.2 Hz, 1H), 2.04 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 159.0, 143.2, 141.5, 139.7, 137.3, 136.8, 131.3, 131.0, 130.0, 129.8, 127.6, 127.4, 126.8, 126.6, 125.1, 82.5, 17.2; IR (film): $\tilde{\upsilon}$ = 3366, 3157, 1651, 1627, 1526, 1464, 1026, 746, 644, 595 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{18}H_{16}N_5O^+$: 318.1349, found: 318.1346.

**2-Amino-1-(2,6-dimethylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6b)**



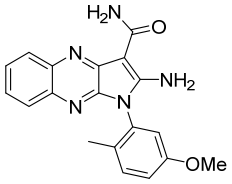
Orange solid; Yield: 4.5 % over two steps; mp 157-159 °C; Purity: 90 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.07 (s, 2H), 7.95 (d, $J$ = 8.3 Hz, 1H), 7.78-7.76 (m, 2H), 7.58 (ddd, $J$ = 8.3, 7.1, 1.4 Hz, 1H), 7.48-7.41 (m, 3H), 7.35-7.34 (m, 2H), 1.96 (s, 6H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 158.6, 142.3, 141.4, 139.7, 137.6, 136.9, 130.0, 129.9, 128.7, 127.6, 126.8, 126.7, 125.2, 82.5, 17.3; IR (film): $\tilde{\upsilon}$ = 3443, 3395, 3309, 1632, 1568, 1523, 1463, 758, 596 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{19}H_{19}N_5O^+$: 332.1506, found: 332.1503.

**2-Amino-1-(2-chloro-6-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6c)**



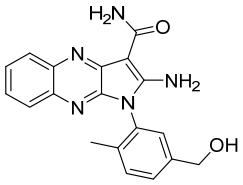
Red solid; Yield: 1 % over two steps; mp 241-243 °C; Purity: 91 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 8.22 (s, 2H), 7.95 (dd, $J$ = 8.3, 1.1 Hz, 1H), 7.79-7.75 (m, 2H), 7.62-7.56 (m, 3H), 7.52 (dd, $J$ = 7.4 Hz, $J$ = 1.2 Hz, 1H), 7.47 (ddd, $J$ = 8.3, 6.9, 1.3 Hz, 1H), 7.39 (s, 1H), 2.08 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.5, 158.5, 142.3, 141.5, 140.7, 139.8, 136.8, 133.7, 131.4, 130.0, 128.7, 127.9, 127.6, 126.8, 125.3, 82.7, 17.7, one C is missing due to overlapping; IR (film): $\tilde{\upsilon}$ = 3359, 3293, 3167, 1649, 1609, 1523, 1455, 1256, 1036, 757, 597 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{18}H_{15}ClN_5O^+$: 352.0960, found: 352.0959.

**2-Amino-1-(5-methoxy-2-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6d)**



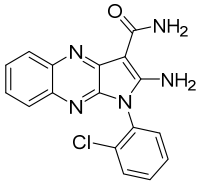
Yellow solid; Yield: 72 %; mp 215-217 °C; $^1$H NMR (300 MHz, DMSO-$d_6$): δ = 7.94 (d, $J$ = 7.7 Hz, 1H), 7.79-7.76 (m, 2H), 7.57 (ddd, $J$ = 8.2, 6.9, 1.3 Hz, 1H), 7.49-7.37 (m, 3H), 7.15-7.09 (m, 2H), 3.78 (s, 3H), 1.94 (s, 3H), $NH_2$-not observed; $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 158.9, 158.2, 143.1, 141.5, 139.6, 136.8, 131.8, 131.6, 128.7, 127.6, 126.8, 126.6, 125.1, 115.9, 115.0, 82.5, 55.4, 16.3; IR (film): ῦ = 3482, 3405, 3295, 3067, 2998, 2930, 1630, 1612, 1507, 1465, 1308, 1227, 1119, 1027, 817, 745, 596 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{19}H_{18}N_5O_2^+$: 348.1455, found: 348.1454.

**2-Amino-1-(5-(hydroxymethyl)-2-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6e)**



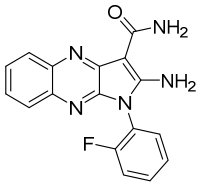
Yellow solid; Yield: 72 %; Purity: 88%; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.03 (s, 2H), 7.94 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.77-7.75 (m, 2H), 7.57 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.49-7.48 (m, 2H), 7.46 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.39-7.36 (m, 2H), 5.34 (t, $J$ = 5.6 Hz, 1H), 4.57 (d, $J$ = 5.6 Hz, 2H), 2.00 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 159.0, 143.1, 142.0, 141.4, 139.6, 136.8, 135.3, 131.0, 130.7, 127.9, 127.6, 127.4, 126.8, 126.6, 125.1, 82.5, 62.0, 16.9; IR (film): ῦ = 3339, 3247, 3162, 1661, 1603, 1148, 1036 ,876, 762, 702 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{19}H_{18}N_5O_2^+$: 348.1455, found: 348.1458.

**2-Amino-1-(2-chlorophenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6f)**



Yellow solid; Yield: 69 %; mp 207-209 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.20 (s, 2H), 7.95 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.80 (dd, $J$ = 8.3, 1.3 Hz, 1H), 7.77-7.75 (m, 3H), 7.69 (td, $J$ = 7.6, 1.7 Hz, 1H), 7.63 (td, $J$ = 7.6, 1.4 Hz, 1H), 7.58 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.46 (ddd, $J$ = 8.3, 7.0, 1.3 Hz, 1H), 7.37 (s, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.5, 158.9, 143.1, 141.5, 139.7, 136.7, 133.2, 132.2, 131.8, 130.5, 129.7, 128.7, 127.5, 126.8, 126.7, 125.2, 82.6; IR (film): ῦ = 3370, 3270, 3173, 1650, 1607, 1522, 1024, 748 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{17}H_{12}ClN_5NaO^+$: 360.0623, found: 360.0622.

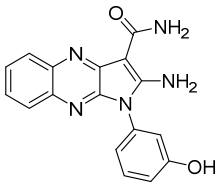
**2-Amino-1-(2-fluorophenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6g)**



Yellow solid; Yield: 55 %; mp 272-275 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.26 (s, 2H), 7.96 (dd, $J$ = 8.3, 1.0 Hz, 1H), 7.78-7.68 (m, 4H), 7.61-7.56 (m, 2H), 7.50-7-46 (m, 2H), 7.39 (s, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.4, 158.9, 158.4 (d, $J$ = 250.2 Hz), 143.2, 141.5, 139.8, 136.6, 132.0 (d, $J$ = 8.0 Hz), 131.6, 127.5, 126.8, 125.6 (d, $J$ = 3.6 Hz), 125.2, 119.5 (d, $J$ = 13.2 Hz), 117.0 (d, $J$ = 19.2 Hz), 82.8, 1C is missing due to overlapping; IR (film): ῦ =

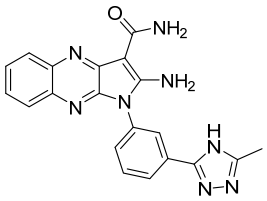
3474, 3394, 3153, 1659, 1507, 1215, 1024, 745 $cm^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{17}H_{12}FN_5NaO^+$: 344.0918, found: 344.0919.

**2-Amino-1-(3-hydroxyphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6h)**



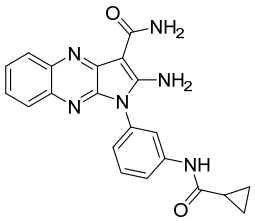
Yellow solid; Yield: 7 %; mp 326-327 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.95 (s, 1H), 8.06 (s, 2H), 7.94 (d, *J* = 8.2 Hz, 1H), 7.80-7.77 (m, 2H), 7.57 (ddd, *J* = 8.3, 7.0, 1.2 Hz, 1H), 7.49-7.43 (m, 2H), 7.35 (s, 1H), 7.00-6.97 (m, 2H), 6.94 (t, *J* = 2.1 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 158.8, 158.3, 143.3, 141.4, 139.6, 136.7, 132.7, 130.5, 127.6, 126.7, 126.6, 125.1, 118.8, 116.3, 115.3, 82.5; IR (film): $\tilde{\upsilon}$ = 3396, 3312, 1635, 1505, 1283, 754 $cm^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{17}H_{13}N_5NaO_2^+$: 342.0962, found: 342.0960.

**2-Amino-1-(3-(5-methyl-4*H*-1,2,4-triazol-3-yl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (6i)**



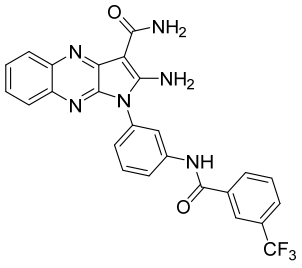
Yellow solid; Yield: 13 %; mp 320-322 °C; Purity: 94 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ =13.83 (s, 1H), 8.20-8.11 (m, 4H), 7.95 (dd, *J* = 8.3, 1.2 Hz, 1H), 7.79-7.73 (m, 3H), 7.62 (d, *J* = 7.6 Hz, 1H), 7.58 (ddd, *J* = 8.3, 6.9, 1.2 Hz, 1H), 7.46 (ddd, *J* = 8.3, 6.9, 1.2 Hz, 1H), 7.37 (s, 1H), 2.41 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 160.0, 158.9, 153.7, 143.5, 141.5, 139.7, 136.7, 133.2, 132.4, 130.2, 128.7, 127.6, 126.8, 126.6, 126.2, 125.6, 125.1, 82.7, 11.5; IR (film): $\tilde{\upsilon}$ = 3370, 3302, 1592, 1524, 1129, 754 $cm^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{20}H_{17}N_8O^+$: 385.1520, found: 385.1522.

**2-Amino-1-(3-(cyclopropanecarboxamido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7a)**



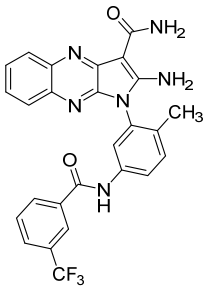
Yellow solid; Yield: 51 %; mp 319-320 °C; Purity: 93 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.51 (s, 1H), 8.09 (s, 2H), 7.94 (dd, *J* = 8.3, 1.1 Hz, 1H), 7.81-7.77 (m, 4H), 7.59-7-56 (m, 2H), 7.47 (ddd, *J* = 8.3, 7.1, 1.3 Hz, 1H), 7.36 (s, 1H), 7.25-7.23 (m, 1H), 1.82 (quint, *J* = 6.2 Hz, 1H), 0.82-0.81 (m, 4H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 171.9, 166.6, 158.8, 143.3, 141.4, 140.6, 139.7, 136.7, 132.2, 130.0, 127.6, 126.8, 126.6, 125.1, 122.8, 119.4, 118.5, 82.5, 14.5, 7.2; IR (film): $\tilde{\upsilon}$ = 3442, 3361, 3136, 1654, 1459, 1240, 1036, 762 $cm^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{21}H_{19}N_6O_2^+$: 387.1564, found: 387.1563.

**2-Amino-1-(3-(3-(trifluoromethyl)benzamido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7b)**



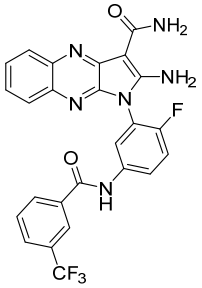
Yellow solid; Yield: 33 %; mp 302-304 °C; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.77 (s, 1H), 8.32 (s, 1H), 8.29 (d, $J$ = 7.9 Hz, 1H), 8.15-8.05 (m, 3H), 8.00-7.95 (m, 3H), 7.83-7.79 (m, 3H), 7.67 (t, $J$ = 8.1 Hz, 1H), 7.60-7.57 (m, 1H), 7.49-7.46 (m, 1H), 7.37-7-36 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 164.1, 158.8, 143.4, 141.4, 140.1, 139.7, 136.7, 135.4, 132.2, 131.8, 130.1, 129.7, 129.1 (q, $J$ = 32.1 Hz), 128.3 (q, $J$ = 3.6 Hz), 127.6, 126.8, 126.6, 125.1, 124.1 (q, $J$ = 3.8 Hz), 124.0, 123.8 (q, $J$ = 272.5 Hz), 120.9, 120.0, 82.6; IR (film): $\tilde{\upsilon}$ = 3452, 3362, 3275, 1639, 1490, 1330, 1171, 757 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{25}H_{18}F_3N_6O_2^+$: 491.1438, found: 491.1443.

**2-Amino-1-(2-methyl-5-(3-(trifluoromethyl)benzamido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7c)**



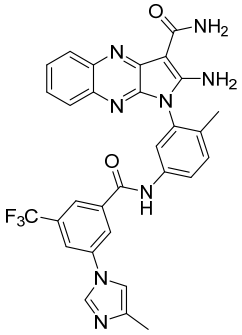
Yellow solid; Yield: 70 %; mp 220-225 °C; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.69 (s, 1H), 8.31 (s, 1H), 8.27 (d, $J$ = 7.9 Hz, 1H), 7.99-7.95 (m, 3H), 7.87 (s, 1H), 7.81-7.77 (m, 3H), 7.58 (t, $J$ = 7.5 Hz, 1H), 7.54 (d, $J$ = 8.4 Hz, 1H), 7.47 (t, $J$ = 7.8 Hz, 1H), 7.39 (s, 1H), 2.01 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 164.0, 158.9, 143.1, 141.5, 139.8, 138.0, 136.9, 135.5, 132.7, 131.9, 131.5, 131.0, 129.8, 129.2 (q, $J$ = 32.1 Hz), 128.3 (q, $J$ = 3.5 Hz), 127.7, 126.9, 126.7, 125.2, 124.2 (q, $J$ = 3.7 Hz), 124.0 (q, $J$ = 272.5 Hz), 121.9, 121.2, 82.61, 16.75; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.07; IR (neat): $\tilde{\upsilon}$ =3301, 3220, 3063, 1637, 1592, 1570, 1509, 1440, 1417, 1332, 1251, 1169, 1121, 1091, 1072, 822, 799, 757, 694, 616, 597, 471, 439, 429 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{26}H_{20}F_3N_6O^+$: 505.1598, found: 505.1594.

**2-Amino-1-(2-fluoro-5-(3-(trifluoromethyl)benzamido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7d)**



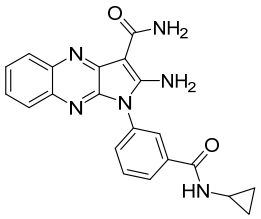
Yellow solid; Yield: 68 %; mp 205-209 °C; Purity: 90 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.80 (s, 1H), 8.33-8.28 (m, 4H), 8.08-8.07 (m, 2H), 7.99 (d, $J$ = 7.5 Hz, 1H), 7.96 (d, $J$ = 8.3 Hz, 1H), 7.83-7.77 (m, 3H), 7.63 -7.58 (m, 2H), 7.49 (t, $J$ = 7.5 Hz, 1H), 7.42 (s, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.5, 164.2, 159.0, 153.8, 143.2, 141.6, 139.9, 136.7, 136.1, 135.3, 131.9, 129.9, 129.4, 127.7, 127.0, 125.4, 124.2, 122.8, 119.4, 117.3, 83.0; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.07, -125.82 (ddd, $J$ = 9.6, 6.4, 4.7 Hz); IR (neat): $\tilde{\upsilon}$ = 3293, 3201, 3073, 1638, 1571, 1462, 1417, 1401, 1332, 1250, 1168, 1122, 1071, 1031, 808, 758, 695, 597, 474, 454, 433 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{17}F_4N_6O_2^+$: 509.1344, found: 509.1343.

**2-Amino-1-(2-methyl-5-((4-methyl-1*H*-imidazol-1-yl)-5-(trifluoromethyl)benzamido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7e)**



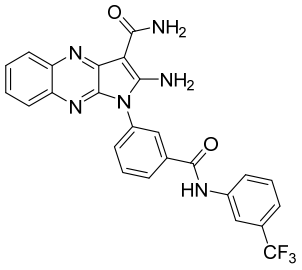
Brown solid; Yield: 90 %; mp 230-235 °C; Purity: 88 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 10.71 (s, 1H), 8.45 (s, 1H), 8.39 (br, 1H), 8.25 (s, 1H), 8.18 (s, 1H), 8.13 (br, 1H), 8.00 (d, $J$ = 8.1 Hz, 1H), 7.96 (d, $J$ = 8.1 Hz, 1H), 7.84 (s, 1H), 7.79 – 7.77 (m, 2H), 7.69 (s, 1H), 7.60 – 7.56 (m, 2H), 7.47 (t, $J$ = 7.7 Hz, 1H), 7.38 (s, 1H), 2.18 (s, 3H), 2.03 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 163.2, 158.9, 143.1, 141.5, 139.8, 139.0, 137.8, 137.3, 136.9, 135.2, 133.0, 131.6, 131.1, 131.0, 130.7, 127.6, 126.9, 126.7, 125.2, 124.5, 122.6, 122.3, 122.0, 121.8, 121.3, 114.2, 82.6, 16.7, 13.5; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.06; IR (neat): ῡ = 3048, 1643, 1571, 1509, 1450, 1376, 1303, 1250, 1174, 1090, 884, 837, 756, 691, 614, 598, 460 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{30}H_{24}F_3N_8O_2^+$: 585.1969, found: 585.1964.

**2-Amino-1-(3-(cyclopropylcarbamoyl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7f)**



Yellow solid; Yield: 46 %; mp 294-296 °C; Purity: 93 %; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 8.54 (d, $J$ = 4.1 Hz, 1H), 8.16 (s, 2H), 8.06 (dt, $J$ = 7.0, 1.9 Hz, 1H), 8.02 (s, 1H), 7.95 (dd, $J$ = 8.3, 1.1 Hz, 1H), 7.78-7.72 (m, 4H), 7.58 (ddd, $J$ = 8.3, 6.9, 1.1 Hz, 1H), 7.47 (ddd, $J$ = 8.3, 6.9, 1.1 Hz, 1H), 7.37 (s, 1H), 2.91-2.86 (m, 1H), 0.73-0.69 (m, 2H), 0.59-0.56 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.5, 166.1, 158.9, 143.4, 141.5, 139.7, 136.7, 135.8, 132.0, 131.5, 129.8, 128.0, 127.6, 127.2, 126.8, 126.7, 125.1, 82.6, 23.0, 5.6; IR (film): ῡ = 3468, 3298, 3117, 1637, 1526, 1457, 1304, 760 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{21}H_{19}N_6O_2^+$: 387.1564, found: 387.1564.

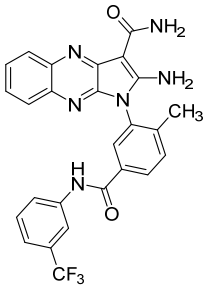
**2-Amino-1-(3-((3-(trifluoromethyl)phenyl)carbamoyl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7g)**



Yellow solid; Yield: 61 %; mp 312-314 °C; $^{1}$H NMR (500 MHz, DMSO-$d_6$): $\delta$ = 10.62 (s, 1H), 8.24-8.22 (m, 5H), 8.08 (d, $J$ = 8.4 Hz, 1H), 7.96 (dd, $J$ = 8.3, 1.0 Hz, 1H), 7.89-7.84 (m, 2H), 7.80-7.78 (m, 2H), 7.64-7.57 (m, 2H), 7.49-7.46 (m, 2H), 7.39 (s, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): $\delta$ = 166.5, 164.6, 158.8, 143.4, 141.5, 139.8, 139.6, 136.7, 135.7, 132.3, 132.2, 130.1, 129.9, 129.3 (q, $J$ = 31.6 Hz), 128.6, 127.7, 127.6, 126.8, 126.7, 125.2, 124.0 (q, $J$ = 272.3 Hz), 123.7, 120.1 (q, $J$ = 3.8 Hz), 116.3 (q, $J$ = 4.0 Hz), 82.7; IR (film): ῡ = 3436, 3272, 1607, 1511,

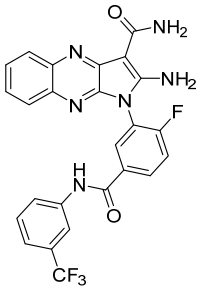
1334, 1112, 751 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{25}H_{18}F_3N_6O_2^+$: 491.1438, found: 491.1440.

**2-Amino-1-(2-methyl-(5-((3-(trifluoromethyl)phenyl)carbamoyl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7h)**



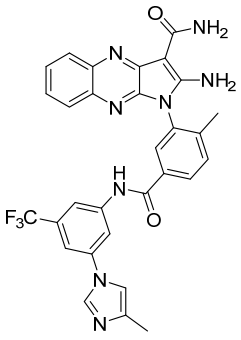
Yellow solid; Yield: 75 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.53 (s, 1H), 8.21-8.15 (m, 5H), 8.08 (d, $J$ = 8.0 Hz, 1H), 7.96 (d, $J$ = 7.4 Hz, 1H), 7.78-7.77 (m, 2H), 7.73 (d, $J$ = 8.1 Hz, 1H), 7.62-7.57 (m, 2H), 7.49-7.45 (m, 2H), 7.40 (s, 1H), 2.13 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 164.4, 159.0, 143.1, 142.0, 141.6, 139.8, 139.8, 136.8, 133.3, 131.7, 131.3, 129.9, 129.5, 129.2, 129.2, 127.7, 126.9, 126.8, 125.3, 123.9, 120.1, 116.5, 82.8, 17.4; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.27; IR (neat): ῡ = 3289, 3213, 2926, 2158, 1638, 1570, 1525, 1492, 1443, 1332, 1167, 1121, 1070, 796, 755, 697, 595, 473, 459, 446, 432 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{26}H_{20}F_3N_6O^+$: 505.1594, found: 505.1594.

**2-Amino-1-(2-fluoro-(5-((3-(trifluoromethyl)phenyl)carbamoyl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7i)**



Yellow solid; Yield: 99 %; Purity: 91 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.64 (s, 1H), 8.39-8.34 (m, 3H), 8.21 (s, 2H), 8.07 (d, $J$ = 7.8 Hz, 1H), 7.97 (d, $J$ = 8.1 Hz, 1H), 7.80-7.74 (m, 3H), 7.64-7.59 (m, 2H), 7.50-7.47 (m, 2H), 7.43 (s, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.4, 163.8, 161.8, 159.8, 158.9, 143.2, 141.6, 139.9, 139.7, 136.7, 132.1, 131.7, 131.5, 130.0, 129.5, 129.3, 127.7, 127.1, 127.0, 125.5, 125.2, 123.9, 123.1, 120.3, 119.8, 119.7, 117.6, 117.4, 116.5, 83.1; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.28, -115.24 – -115.33 (m); IR (neat): ῡ =3292, 3207, 3070, 1641, 1571, 1514, 1491, 1444, 1332, 1275, 1167, 1116, 1094, 1070, 797, 754, 697, 614, 597, 481, 419 cm$^{-1}$. HRMS (ESI), *m/z*: calcd for $C_{25}H_{17}F_4N_6O_2^+$: 509.1344, found: 509.1343.

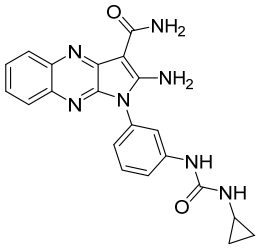
**2-Amino-1-(2-methyl-5-((3-(4-methyl-1*H*-imidazol-1-yl)-(5-(trifluoromethyl)phenyl)carbamoyl)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7j)**



Orange solid; Yield: 85 %; mp 233-238 °C; Purity: 91 % $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.66 (s, 1H), 8.27 (s, 1H), 8.22 – 8.15 (m, 5H), 7.97 (d, $J$ = 8.4 Hz, 1H), 7.78 – 7.73 (m, 4H), 7.59 (t, $J$ = 7.6 Hz, 1H), 7.49 – 7.46 (m, 2H), 7.40 (s, 1H), 7.09 (br, 2H), 2.16 (s, 3H), 2.15 (s, 3H); $^{13}$C-NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 164.5, 159.0, 143.2, 142.4, 141.6, 141.1, 139.8, 138.9, 138.0, 136.8, 135.0, 133.0, 131.8, 131.4,

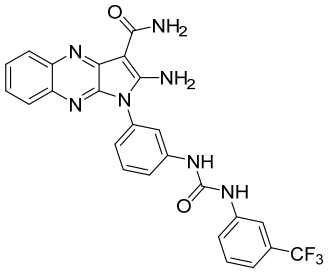
129.6, 129.2, 127.7, 126.9, 126.9, 125.3, 124.7, 115.0, 114.3, 111.9, 109.5, 109.5, 82.8, 17.5, 13.6; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.34; IR (neat): ῦ = 3303, 2923, 2202, 1683, 1653, 1614, 1565, 1499, 1478, 1444, 1325, 1309, 1268, 1241, 1226, 1193, 1170, 1123, 1103, 1082, 1021, 1008, 929, 863, 807, 758, 750, 694, 670, 627, 597, 579, 523, 470, 452, 439, 411 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{30}H_{24}F_3N_8O_2^+$: 585.1969, found: 585.1970.

**2-Amino-1-(3-(3-cyclopropylureido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7k)**



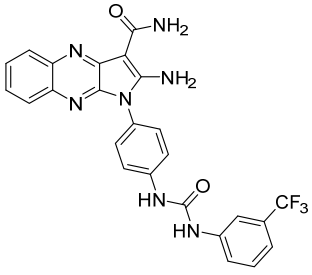
Yellow solid; Yield: 0.6 % over 2 steps; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.64 (s, 1H), 8.06 (s, 2H), 7.94 (dd, *J* = 8.3, 1.2 Hz, 1H), 7.79-7.77 (m, 2H), 7.63-7.62 (m, 2H), 7.58 (ddd, *J* = 8.3, 6.9, 1.3 Hz, 1H), 7.51-7.45 (m, 2H), 7.35 (s, 1H), 7.09 (d, *J* = 8.1 Hz, 1H), 6.53 (s, 1H), 2.56-2.52 (m, 1H), 0.65-0.61 (m, 2H), 0.42-0.39 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 158.8, 155.8, 143.4, 141.7, 141.4, 139.7, 136.7, 132.1, 129.9, 127.6, 126.8, 126.6, 125.1, 120.9, 118.4, 117.4, 82.5, 22.3, 6.3; IR (film): ῦ = 3411, 3289, 1639, 1514, 1292, 1024, 761 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{21}H_{19}N_7NaO_2^+$: 424.1492, found: 424.1494.

**2-Amino-1-(3-(3-(3-(trifluoromethyl)phenyl)ureido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7l)**



Yellow solid; Yield: 15 %; mp 303-305 °C; Purity: 91 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 9.45 (s, 1H), 9.40 (s, 1H), 8.12 (s, 2H), 8.02 (s, 1H), 7.95 (d, *J* = 7.5 Hz, 1H), 7.80-7.78 (m, 2H), 7.72-7.68 (m, 2H), 7.60-7.45 (m, 5H), 7.36 (s, 1H), 7.31 (d, *J* = 7.6 Hz, 1H), 7.20 (d, *J* = 8.4 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 158.8, 152.4, 143.4, 141.4, 140.7, 140.3, 139.7, 136.7, 132.3, 130.1, 129.8, 129.4 (q, *J* = 31.4 Hz), 127.6, 126.8, 126.6, 125.1, 124.1 (q, *J* = 31.4 Hz), 121.9, 121.8, 118.9, 118.1 (q, *J* = 31.4 Hz), 118.0, 114.0 (q, *J* = 31.4 Hz), 82.6; IR (film): ῦ = 3504, 3405, 3292, 1637, 1601, 1566, 1331, 1108, 754 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{25}H_{19}F_3N_7O_2^+$: 506.1547, found: 506.1547.

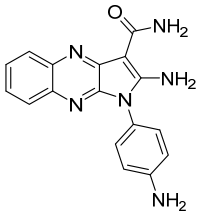
**2-Amino-1-(4-(3-(3-(trifluoromethyl)phenyl)ureido)phenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7m)**



Yellow solid; Yield: 99 %; mp 205-210 °C; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 9.61 (br, 1H), 8.07 (s, 1H), 7.92 (d, *J* = 7.8 Hz, 1H), 7.78-7.74 (m, 4H), 7.68 (d, *J* = 7.5 Hz, 1H), 7.57-7.51 (m, 2H), 7.49-7.44 (m, 3H), 7.32 (s, 1H), 7.30 (s, 1H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ = 166.7, 159.1, 152.7, 143.6, 141.4, 140.8, 140.6, 139.7, 136.8, 129.8, 129.7, 129.3, 129.1, 127.5, 126.8, 126.5,

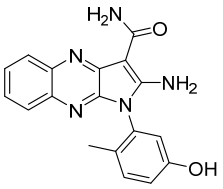
125.6, 125.5, 125.0, 122.9, 121.9, 119.3, 118.0, 114.2, 82.6; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.24; IR (neat): ῦ = 3305, 1637, 1567, 1515, 1493, 1446, 1336, 1314, 1232, 1174, 1120, 1096, 1069, 794, 756, 698, 662, 614, 597, 528, 516, 473, 457, 421, 413 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{19}F_3N_7O_2^+$: 506.1547, found: 506.1549.

**2-Amino-1-(*N*-Boc-*p*-phenylamine)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (7n)**



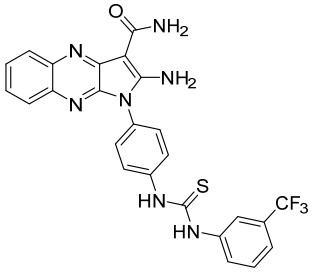
Yellow solid; Yield: 86 %; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 7.92 (d, *J* = 7.8 Hz, 1H), 7.77-7.70 (m, 2H), 7.56 (t, *J* = 7.5 Hz, 1H), 7.46 (t, *J* = 7.3 Hz, 1H), 7.30 (br, 1H), 7.19 (d, *J* = 8.6 Hz, 2H), 6.81 (d, *J* = 8.5 Hz, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.7, 159.6, 144.0, 141.4, 139.6, 136.9, 129.4, 127.6, 126.7, 126.5, 125.1, 115.0, 82.4, 2C are missing due to overlapping; IR (neat): ῦ = 3337, 3186, 2957, 1635, 1567, 1519, 1444, 1414, 1365, 1315, 1297, 1265, 1237, 1193, 1167, 1058, 1030, 832, 797, 753, 707, 612, 596, 518, 474, 461, 447, 433, 421 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{17}H_{15}N_6O^+$: 319.1302, found: 319.1306.

**2-Amino-1-(5-hydroxy-2-methylphenyl)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (11d)**



To a solution of 2-amino-1-(5-methoxy-2-methylphenyl)-1H-pyrrolo[2,3-b]quinoxaline-3-carboxamide (**6d**, 135 mg, 0.389 mmol) in anhydrous $CH_2Cl_2$ (5.00 mL) at -78 °C $BBr_3$ (1 M in $CH_2Cl_2$, 972 µL, 0.972 mmol) was added. The reaction was allowed to warm to room temperature and was then heated for 4 h at 130 °C. The formed solid was filtered off and purified by flash chromatography on silica gel (gradient EtOAc:MeOH 99:1 to 98:2) to afford the desired product in pure form (88.2 mg, 68 % yield). mp 272-274 °C; Purity: 94 %; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.73 (s, 1H), 8.02 (s, 2H), 7.94 (dd, *J* = 8.3, 1.3 Hz, 1H), 7.78 (dd, *J* = 8.2, 1.3 Hz, 1H), 7.75 (s, 1H), 7.57 (ddd, *J* = 8.3, 6.9, 1.3 Hz, 1H), 7.46 (ddd, *J* = 8.2, 6.9, 1.3 Hz, 1H), 7.37 (s, 1H), 7.30 (d, *J* = 8.4 Hz, 1H), 6.94 (dd, *J* = 8.4, 2.5 Hz, 1H), 6.81 (d, *J* = 2.5 Hz, 1H), 1.88 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.6, 158.8, 156.3, 143.0, 141.4, 139.6, 136.8, 131.9, 131.3, 127.6, 126.8, 126.6, 125.1, 117.2, 116.2, 82.4, 16.2, one C is missing due to overlapping; IR (film): ῦ = 3455, 3380, 3292, 3200, 1633, 1511, 1461, 1021, 753, 595 cm$^{-1}$; HRMS (ESI): *m/z*: calcd for $C_{18}H_{16}N_5O_2^+$: 334.1299, found: 334.1296.

**2-Amino-1-(4-(3-(3-(trifluoromethyl)phenyl)thioureido)phenyl)-1*H*-pyrrolo[2,3-b]quinoxaline-3-carboxamide (12n)**



To a solution of 2-amino-1-(*N*-Boc-*p*-phenylamine)-1*H*-pyrrolo[2,3-*b*]quinoxaline-3-carboxamide (**7n**, 10 mg, 0.031 mmol) in DMF (0.10 mL) 3-(trifluoromethyl)phenyl isothiocyanate (3.7 µL, 0.024 mmol) was added. The reaction

mixture was stirred at 25 °C for 12 hours. The resulting precipitate was filtered off and washed with DCM and hexane, affording the desired thiourea as a yellow solid in pure form (8.63 mg, 69 % yield). $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 10.22 (br, 2H), 8.08 (br, 2H), 8.02 (s, 1H), 7.95 (dd, $J$ = 8.2, 1.3 Hz, 1H), 7.86-7.73 (m, 5H), 7.64-7.53 (m, 4H), 7.52-7.43 (m, 2H), 7.35 (s, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 166.7, 159.1, 143.6, 141.5, 140.4, 140.0, 139.8, 136.8, 129.6, 129.2, 129.0, 128.8, 127.9, 127.7, 127.3, 126.9, 126.7, 125.2, 124.2 123.0, 120.8, 119.7, 82.7; $^{19}$F NMR (282 MHz, DMSO-$d_6$): δ = -61.16; IR (neat): ῡ = 3302, 3193, 2917, 2848, 1639, 1602, 1515, 1448, 1330, 1164, 1120, 1095, 1069, 797, 758, 696, 667, 596, 521, 475, 464, 449, 439 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{19}F_3N_7OS^+$: 522.1318, found: 522.132.

### 2.3.2 Differential Scanning Fluorimetry

To a 2 µM solution of EphA3 in 10 mM Tris HCl buffer (pH 8.0) containing 100 mM NaCl, 5X SPYRO Orange fluorescent dye together with the test compounds prediluted in DMSO at a concentration of 20 µM (1.4 % final DMSO concentration) were added. A final volume of 40 µL was added per well in a 96 well plate and incubated at room temperature for 1h. The plates were centrifuged for 3 min at 450 rcf. The temperature gradient was performed in the range of 25-75 °C (50 acquisitions per degree at a ramp rate of 4.4 °C/s) using a LightCycler 480 Real-Time PCR System. The detection of protein unfolding was performed with an excitation wavelength of 488 nm and an emission of 580 nm. All measurements were performed in triplicate.

### 2.3.3 FRET based enzymatic assay

Selected compounds were tested in the Z'-LYTE$^{TM}$ Kinase Assay Kit–Tyr 1 Peptide (Invitrogen, USA) in a Corning 384 well microtiter plate. Fluorescence progress curves were measured upon excitation at 400 nm and emission at 445 and 520 nm. The assay contained a final concentration of EphB4 and ATP of 25 ng/µL and 125 µM (which is near its $K_m$), respectively, and was run at room temperature for 2 h. $IC_{50}$ values (inhibitor concentration at which enzyme activity is reduced by 50%) are determined after carrying out assays at ten different concentrations between 20 µM and 10 pM.

### 2.3.4 Surface Plasmon Resonance Measurements

The affinity of **7m** and **11d** towards EphA3 was determined using Surface Plasmon Resonance measurements on a ProteOn XPR36 (Bio-Rad Laboratories, AG). EphA3 was immobilized on a GLH chip (Bio-Rad Laboratories AG) by standard amine coupling procedures achieving a RU of 4000-6000 and equilibrated with steriled-filtered buffer (10mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.005 % Tween 20 pH 8.0) with a flow rate of 60 µL/min. Association of **7m** and **11d** was performed at 25 °C using serial dilutions of

the inhibitors in duplicates for 400 s, whereas dissociation of the inhibitors from the chip was monitored for 36,000 seconds. Data were normalized using interspot referencing and subtraction of a separate analyte channel run with buffer and BSA. All sensograms were fitted using a 1:1 Langmuir model provided by the ProteOn Manager Software (Bio-Rad Laboratories AG), and the association ($k_{on}$) and dissociation ($k_{off}$) rate constants were used to determine the equilibrium dissociation constants ($K_D$).

### 2.3.5 X-Ray Diffraction Analysis

#### I. Protein expression and purification

A clone of the EphA3 kinase domain (residues: 606-947) was obtained from Prof. Sirano Dhe-Paganon's group[120] and expressed in *Escherichia coli* strain BL21 (DE3). Cells expressing EphA3 were induced with a 1 mM solution of isopropyl-$\beta$-D-thiogalactopyranoside (IPTG) for 12 h at 15 °C. Cell pellets were resuspended in buffer A (50 mM Tris, pH 8.0, and 100 mM NaCl, supplemented with protease inhibitors) and lysed by sonication. After centrifugation at 15,000 rpm for 1 h, the soluble fraction of EphA3 was purified using HisTrap FF crude and HiTrap Q HP columns (GE Healthcare), followed by gel filtration chromatography (Superdex75; GE Healthcare). The appropriate fractions were combined and concentrated to ~10 mg/mL using Amicon filter devices (10 kDa as cutoff) in a storage solution (100 mM sodium chloride and 10 mM Tris-HCl pH 8.0, 5% glycerol). The resulting solution was aliquoted and stored at -80 °C for further usage.

#### II. Crystallization, Data Collection, and Structure Determination

Apo crystals of the EphA3 kinase domain were grown at 20 °C using the hanging drop vapor diffusion method. Equal volumes of protein and reservoir solutions (0.1 M sodium cacodylate pH 6.5, 0.15 M ammonium sulfate, 22.5 % PEG 3350) were mixed and crystals appeared after 1 to 2 days. A 5 mM solution of inhibitor (in 100 % DMSO) was added into the hanging drop to reach a final DMSO concentration of 10 % (v/v). The crystals were soaked with compounds **A** or **B** for 1 to 24 h and flash-frozen in liquid nitrogen without extra cryoprotectant. Cocrystallization trials were performed for compound **7m**. EphA3 was first mixed with a 5 mM compound **7m** solution (in 100 % DMSO) to a final 10 % DMSO concentration and then incubated on ice for 30 minutes. EphA3-**7m** crystals were grown under the same conditions as described above for the EphA3 apo crystals. Data sets were collected on a Pilatus 6M detector at the Swiss Light Source beamline X06SA of the Paul Scherrer Institute (Villigen, Switzerland) and indexed, integrated and scaled with the XDS[121] and CCP4 programs.[122] The structures were solved by molecular replacement with PHASER[123] using the apo EphA3 kinase domain structure (PDB entry 2GSF) as a search model and refined with PHENIX.[124] The atomic coordinates and structure factors of EphA3

in complex with the inhibitors **A**, **B**, **7m** have been deposited with the Protein Data Bank as entries 4P4C, 4P5Q and 4P5Z respectively.

| | **Compound A** | **Compound B** | **Compound 7m** |
|---|---|---|---|
| **Space group** | P 1 21 1 | P 1 21 1 | P 1 21 1 |
| **Unit cell** | | | |
| **a (Å)** | 53.91 | 53.37 | 52.81 |
| **b (Å)** | 38.16 | 38.2 | 38.24 |
| **c (Å)** | 75.53 | 75.77 | 75.46 |
| **Resolution range (Å)** | 38.16-1.6 | 47.48-1.35 | 47.02-2.0 |
| **Unique reflections** | 39425(5697) | 64017(2998) | 20066(2819) |
| **<I/σ(I)>** | 13.9(3.7) | 15.9(2.1) | 15.6(3.0) |
| **R merge** | 0.052 (0.317) | 0.049(0.705) | 0.069(0.586) |
| **Completeness (%)** | 98.2 (97.6) | 97.1(93.2) | 99.1(96.2) |
| **Multiplicity** | 3.5(3.5) | 4.1(4.0) | 6.5(6.5) |
| **Refinement** | | | |
| **Resolution range (Å)** | 33.64-1.60 | 39.19-1.35 | 47.02-2.00 |
| **R factor/R free** | 17.78/20.14 | 15.70/18.19 | 17.96/21.86 |
| **Mean B factors (A2)** | 18.90 | 24.40 | 43.90 |
| **RMS bonds (Å)** | 0.0087 | 0.0055 | 0.0151 |
| **RMS angles (°)** | 1.343 | 1.258 | 1.724 |

### 2.3.6 Selectivity profile

The selectivity profiling was performed at DiscoveRx in a library of 456 kinases at a concentration of 1 µM. Briefly, kinases were mostly expressed as fusion proteins to T7 phase and grown in 24-well blocks in *E. Coli* (derived from the BL21 strain). The rest of the kinases were expressed in HEK-293 cells and tagged with DNA for qPCR detection. Affinity resins were generated by mixing streptavidin-coated magnetic beads with biotinylated small molecule ligands for 30 minutes at 25 °C, followed by biotin and blocking buffer addition (SeaBlock (Pierce), 1 % BSA, 0.05 % Tween 20, 1 mM DTT). Inhibitors were kept at 40x stocks in 100 % DMSO and added to the corresponding 384-well plates (40 µL). After 1 hour of incubation at 25 °C while shacking, affinity beads were washed with 0.05% Tween 20 in PBS. The beads were re-suspended in PBS buffer containing 0.05 % Tween 20 and 0.5 µM non-biotinylated affinity ligand, to then be incubated at 25 °C for 30 minutes while shacking. The kinase concentration present in the eluate was determined by qPCR.

### 2.3.7 Cellular phosphorylation assays

Cellular phosphorylation experiments were performed at ProQinase GmbH. Mouse embryonic fibroblast cells, stably transfected to overexpress full-length human EphB4, were plated at $4\times10^4$ cells/well in DMEM (Dulbelcco's modified Eagle's medium) supplemented with 10 % FCS in 48-well culture dishes. Medium was replaced by DMEM without FCS before test compounds prediluted in 100 % DMSO were added (final DMSO concentration of 1 %). After incubation for 90 min at 37 °C, cells were stimulated for 2 h at 4 °C using murine ephrinB2-Fc at a final concentration of 2 µg/ml. Quantification of EphB4 phosphorylation was assessed in a 96-well plate via sandwich ELISA using a myc capture antibody and an antiphosphotyrosine detection antibody.

### 2.3.8 *In house* cell culture and $GI_{50}$ determination

MDA-MB-231 cells obtained from the UZH Cancer Institute and A498 cells obtained from Prof. Krek (Institute of Cell Biology, ETH) were cultured in DMEM supplemented with 10 % (v/v) fetal bovine serum. K562 cells obtained from Dr. Silvio Hemmi (Institute of Molecular Life Sciences, UZH), KM12 (UZH Cancer Institute) and HOP-92 (purchased from the NCI) were cultured using RPMI medium supplemented with 10 % (v/v) fetal bovine serum. All the media were additionally supplemented with 100 units/mL of penicillin, 100 µg/mL of streptomycin, 4.5 g/L glucose, 0.11g/L sodium pyruvate and 2mM glutamine and the cells were grown at 37 ºC in 5 % $CO_2$ atmosphere with 80 % relative humidity.

MDA-MB-231, A498, KM12 and HOP-92 cells were plated at 10,000 cells per well (100 µL per well) in 96-well culture dishes and allowed to incubate for 24 h. In the case of HT29, 20,000 cells per well were seeded and incubated for 24 h. The old media was removed, cells were washed with PBS (phosphate-buffered saline) and fresh media was added. A 5 mM solution of inhibitor (in 100% DMSO) was serially diluted in the culture media (12 different concentrations were used) and allowed to incubate for 72 h (MDA-MB-231) or 48h (A498, KM12, HT29 and HOP-92). Control cells were treated with the same DMSO concentrations. After the incubation period the media was once more removed and the cells were washed with PBS to then be incubated with fresh media containing 86 nM resazurin. After 4 hours, the fluorescence was quantified using a fluorescence microplate reader (Biotek, FLx800$^{TM}$) at the respective excitation and emission wavelength of 560 and 590 nm. Resazurin becomes fluorescent under mitochondrial reduction and its emission directly correlates with the metabolic viability of the cells. The measured fluorescence values were corrected from the control samples containing DMSO and $IC_{50}$ values were determined by plotting the fraction of metabolically active cells against the log of the drug concentration. All measurements were performed at least in triplicate.

In the case of K-562 leukemia cells, they were seeded at a density of 20,000 cells per well in 100µL of RMPI media in a 96 well microtiter plates. After 24 hours, 12.5 µL of a 10 fold concentrated drug or DMSO solution in RMPI media was added in every well. After 48 hour incubation, cell viability was studied by measuring the ability of the cells to process resazurin. Resazurin was added to every well to obtain a final concentration of 86 nM, and after 3 hours, the fluorescence was quantified using a fluorescence microplate reader (Biotek, FLx800$^{TM}$) at the respective excitation and emission wavelength of 560 and 590 nm as described above. All measurements were performed at least in triplicate.

### 2.3.9 Antiproliferative activity against patient derived tumor cell lines

The assessment of antiproliferative activity against patient derived tumor cell lines was performed in Oncotest. Cell lines were routinely passaged once or twice weekly and maintained in culture for up to 20 passages. All cells were cultured in RPMI 1640 medium supplemented with 10 % (v/v) fetal calf serum and 0.1 mg/mL gentamicin (medium and all components from PAA, Cölbe, Germany) at 37 °C in a humidified atmosphere with 5 % $CO_2$. A modified propidium iodide (PI) assay was used to assess the anticancer activity of the compounds. Briefly, cells were harvested from exponential phase cultures, counted, and plated in 96-well flat-bottom microtiter plates at a cell density of 4000−20 000 cells per well. After a 24 h recovery period, allowing the cells to resume exponential growth, 10 µL of culture medium (four control wells/plate) and culture medium with the test compound were added by the liquid handling robotic system and treatment was continued for four days. The compounds were applied in half log increments at 10 concentrations in duplicate. Next, cells were washed with 200 µL of PBS to remove dead cells. Subsequently, 200 µL of a solution containing 7 µg/mL propidium iodide (PI) and 0.1% (v/v) Triton X-100 was added. After an incubation period of 1−2 h at room temperature, fluorescence (FU) was measured using the Cytofluor 4000 microplate reader (excitation $\lambda$ = 530 nm, emission $\lambda$ = 620 nm) to quantify the amount of attached viable cells. For calculations, the mean value of duplicate/quadruplicate (untreated control) data was used. Quality criteria for a successful assay included fluorescence intensity signal of >500 units from the untreated control wells.

### 2.3.10 Spheroid-based cellular angiogenesis assay

The experiments were performed in ProQinase Gmbh and pursued in modification of the originally published protocol.[125] In brief, spheroids were prepared as described[126] by pipetting 500 HUVEC in a hanging drop on plastic dishes to allow overnight spheroid aggregation. 50 HUVEC spheroids were then seeded in 0.9 ml of a collagen gel and pipetted into individual wells of a 24 well plate to allow polymerization. The test

compounds in combination with the respective growth were added after 30 min by pipetting 100 µl of a 10-fold concentrated working solution on top of the polymerized gel. Plates were incubated at 37 °C for 24 hours and fixed by adding 4 % Roti-Histofix (Roth, Karlsruhe, Germany).

Sprouting intensity of HUVEC spheroids treated with the growth factors and inhibitors were quantitated by an image analysis system determining the cumulative sprout length per spheroid (CSL) using an inverted microscope and the digital imaging software Analysis 3.2 (Soft imaging system, Münster, Germany). The mean of the cumulative sprout length of 10 randomly selected spheroids was analyzed as an individual data point. The reduction of sprouting below negative control levels could be due to neutralization of basal concentrations of the respective target or the induction of cell death. With regard to the visual characterization cell death can neither be confirmed nor excluded.

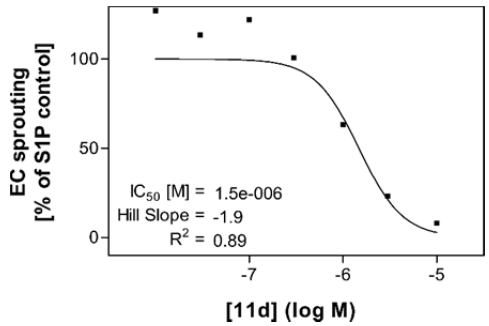



**Figure 19.** HUVEC spheroids were embedded in a 3D collagen gel, stimulated with VEGF-A (25 ng/ml) and treated for 24 h with different concentrations of **11d** (in M, as indicated). The cumulative sprout length of 10 randomly selected spheroids per data point was analyzed and the relative inhibition by the test compound determined. Fitting of inhibition curves and calculation of $IC_{50}$ values was performed with GraphPad Prism 5.

## 2.3.11 *In Vivo* evaluation

### I. Pharmacokinetic Evaluation in Mice

Compounds in Table 4 were administered via oral (PO) and intravenous (IV, tail vein) routes at doses of, 5 mg/kg and 1 mg/kg, respectively, in formulations consisting of, for IV dosing, each compound within a clear solution of DMSO/Solutol® HS 15/phosphate-buffered saline, pH 7.4 (PBS) (5/5/90, v/v/v), and of, for PO dosing, each compound as a heterogeneous suspension in DMSO/1% methylcellulose (5/95, v/v). Blood samples collected over an 8-hour time period (at 3, 10, 30, 60, 120, 240, 360 and 1440 minutes, and 10, 30, 60, 120, 240, 360, 480 and 1440 minutes, for IV and PO administrations, respectively) following administration were centrifuged at 2500 G for 15 minutes at 4 °C within one hour of collection and stored at -20 °C until analysis.

Compound concentration in plasma at each time point was then determined following acetonitrile precipitation of each sample followed by HPLC-MS analysis performed in triplicates, i.e., based on the mean of the compound concentration at that time point in blood samples of three mice. Pharmacokinetic data was generated by non-compartmental analysis (NCA) of the plasma data using WinNonLin. Bioavailability (F) was calculated as a ratio of areas under the curve (AUClast) of the plot of compound concentration versus time up to the time point of the last observation that was greater than the limit of quantification. Animal studies of pharmacokinetics were performed in compliance with the recommendations of the *Guide for Care and Use of Laboratory Animals* with respect to restraint, husbandry, surgical procedures, feed and fluid regulation, and veterinary care in a facility that was monitored and inspected by Cerep Institutional Animal Care and Use Committee.

**II. Xenograft mouse model efficacy evaluation**

Xenograft experiments were conducted on 9 week old 18-27 g female athymic NCr Nu/Nu mice (Charles River Laboratories, Inc.), injected at the flank with $5x10^6$ MDA-MB-231 cells. Tumor size was measured by digital caliper. Once tumor volume reached 108-126 $mm^3$, mice were sorted into a treatment and vehicle-control cohort consisting of 9 animals each such that the group mean tumor volumes were of 112 $mm^3$. The treatment cohort received oral doses corresponding to 50 mg/kg of compound **11d** twice daily for 21 days (with the exception of a single dose on the first day), formulated as per the PO formulation of the pharmacokinetic study. The control group received the same formulation but excluding compound **11d**, according to the same protocol. Following treatment initiation, body weight and tumor volume of all mice in both groups was evaluated bi-weekly up to day 19. The %TGI was calculated according to the formula: $[1-(MTV_{compound\ treated}/MTV_{control})]$ x 100, where MTV corresponds to the Median Tumor Volume for the group. Animal studies of efficacy were performed in compliance with the recommendations of the *Guide for Care and Use of Laboratory Animals* with respect to restraint, husbandry, surgical procedures, feed and fluid regulation, and veterinary care by CRL-DRS-NC in an AAALAC accredited facility.

## 2.4 References

1. Matthews, D. J.; Gerritsen, M. E. *Targeting Protein Kinases for Cancer Therapy.* John Wiley & Sons: United States of America, 2010.

2. Zhang, Z. Y. Protein tyrosine phosphatases: Structure and function, substrate specificity, and inhibitor development. *Annu. Rev. Pharmacol. Toxicol.* **2002,** 42, 209-234.

3. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002,** 298, 1912-34.
4. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*. Garland Science, Taylor & Francis Group: United States of America, 2002.
5. Johnson, L. N.; Lowe, E. D.; Noble, M. E. M.; Owen, D. J. The structural basis for substrate recognition and control by protein kinases. *FEBS Lett.* **1998,** 430, 1-11.
6. Huse, M.; Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **2002,** 109, 275-282.
7. Zhang, J.; Yang, P. L.; Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* **2009,** 9, 28-39.
8. Lafleur, K. Design and Synthesis of Selective Kinase Inhibitors. University of Zürich, Zürich, 2011.
9. Chico, L. K.; Van Eldik, L. J.; Watterson, D. M. Targeting protein kinases in central nervous system disorders. *Nat. Rev. Drug Discov.* **2009,** 8, 892-909.
10. Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat. Rev. Drug Discov.* **2002,** 1, 309-15.
11. Kumar, R.; Singh, V. P.; Baker, K. M. Kinase inhibitors for cardiovascular disease. *J. Mol. Cell. Cardiol.* **2007,** 42, 1-11.
12. Doerig, C.; Billker, O.; Pratt, D.; Endicott, J. Protein kinases as targets for antimalarial intervention: Kinomics, structure-based design, transmission-blockade, and targeting host cell enzymes. *Bba-Proteins Proteom.* **2005,** 1754, 132-150.
13. Tamaoki, T.; Nomoto, H.; Takahashi, I.; Kato, Y.; Morimoto, M.; Tomita, F. Staurosporine, a Potent Inhibitor of Phospholipid/Ca++Dependent Protein-Kinase. *Biochem. Biophys. Res. Commun.* **1986,** 135, 397-402.
14. http://www.brimr.org/PKI/PKIs.htm.
15. Zuccotto, F.; Ardini, E.; Casale, E.; Angiolini, M. Through the "gatekeeper door": exploiting the active kinase conformation. *J. Med. Chem.* **2010,** 53, 2681-94.
16. Liu, Y.; Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* **2006,** 2, 358-64.
17. Backes, A. C.; Zech, B.; Felber, B.; Klebl, B.; Muller, G. Small-molecule inhibitors binding to protein kinases. Part I: exceptions from the traditional pharmacophore approach of type I inhibition. *Expert Opin. Drug Dis.* **2008,** 3, 1409-1425.
18. Davis, M. I.; Hunt, J. P.; Herrgard, S.; Ciceri, P.; Wodicka, L. M.; Pallares, G.; Hocker, M.; Treiber, D. K.; Zarrinkar, P. P. Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2011,** 29, 1046-51.
19. Knight, Z. A.; Shokat, K. M. Features of selective kinase inhibitors. *Chem. Biol.* **2005,** 12, 621-637.
20. Lombardo, L. J.; Lee, F. Y.; Chen, P.; Norris, D.; Barrish, J. C.; Behnia, K.; Castaneda, S.; Cornelius, L. A. M.; Das, J.; Doweyko, A. M.; Fairchild, C.; Hunt, J. T.; Inigo, I.; Johnston, K.; Kamath, A.; Kan, D.; Klei, H.; Marathe, P.; Pang, S. H.; Peterson, R.; Pitt, S.; Schieven, G. L.; Schmidt, R. J.; Tokarski, J.; Wen, M. L.; Wityak, J.; Borzilleri, R. M. Discovery of N-(2-chloro-6-methylphenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual

Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J. Med. Chem.* **2004,** 47, 6658-6661.

21. Kornev, A. P.; Haste, N. M.; Taylor, S. S.; Ten Eyck, L. F. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** 103, 17783-17788.

22. Kornev, A. P.; Taylor, S. S.; Ten Eyck, L. F. A helix scaffold for the assembly of active protein kinases. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** 105, 14377-14382.

23. Morphy, R. Selectively nonselective kinase inhibition: striking the right balance. *J. Med. Chem.* **2010,** 53, 1413-37.

24. Pargellis, C.; Tong, L.; Churchill, L.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Grob, P. M.; Hickey, E. R.; Moss, N.; Pav, S.; Regan, J. Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat. Struct. Biol.* **2002,** 9, 268-72.

25. Regan, J.; Pargellis, C. A.; Cirillo, P. F.; Gilmore, T.; Hickey, E. R.; Peet, G. W.; Proto, A.; Swinamer, A.; Moss, N. The kinetics of binding to p38 MAP kinase by analogues of BIRB 796. *Bioorg. Med. Chem. Lett.* **2003,** 13, 3101-3104.

26. Angell, R. M.; Angell, T. D.; Bamborough, P.; Brown, D.; Brown, M.; Buckton, J. B.; Cockerill, S. G.; Edwards, C. D.; Jones, K. L.; Longstaff, T.; Smee, P. A.; Smith, K. J.; Somers, D. O.; Walker, A. L.; Willson, M. Biphenyl amide p38 kinase inhibitors 2: Optimisation and SAR. *Bioorg. Med. Chem. Lett.* **2008,** 18, 324-8.

27. Weiss, M. M.; Harmange, J. C.; Polverino, A. J.; Bauer, D.; Berry, L.; Berry, V.; Borg, G.; Bready, J.; Chen, D.; Choquette, D.; Coxon, A.; DeMelfi, T.; Doerr, N.; Estrada, J.; Flynn, J.; Graceffa, R. F.; Harriman, S. P.; Kaufman, S.; La, D. S.; Long, A.; Neervannan, S.; Patel, V. F.; Potashman, M.; Regal, K.; Roveto, P. M.; Schrag, M. L.; Starnes, C.; Tasker, A.; Teffera, Y.; Whittington, D. A.; Zanon, R. Evaluation of a series of naphthamides as potent, orally active vascular endothelial growth factor receptor-2 tyrosine kinase inhibitors. *J. Med. Chem.* **2008,** 51, 1668-1680.

28. Shah, N. P.; Nicoll, J. M.; Nagar, B.; Gorre, M. E.; Paquette, R. L.; Kuriyan, J.; Sawyers, C. L. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* **2002,** 2, 117-125.

29. Capdeville, R.; Buchdunger, E.; Zimmermann, J.; Matter, A. Glivec (ST1571, Imatinib), a rationally developed, targeted anticancer drug. *Nat. Rev. Drug Discov.* **2002,** 1, 493-502.

30. Blanc, J.; Geney, R.; Menet, C. Type II kinase inhibitors: an opportunity in cancer for rational design. *Anticancer Agents Med. Chem.* **2013,** 13, 731-47.

31. Lafleur, K.; Dong, J.; Huang, D.; Caflisch, A.; Nevado, C. Optimization of inhibitors of the tyrosine kinase EphB4. 2. Cellular potency improvement and binding mode validation by X-ray crystallography. *J. Med. Chem.* **2013,** 56, 84-96.

32. Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J. Med. Chem.* **2009,** 52, 6433-46.

33. Zhao, H. T.; Dong, J.; Lafleur, K.; Nevado, C.; Caflisch, A. Discovery of a Novel Chemotype of Tyrosine Kinase Inhibitors by Fragment-Based Docking and Molecular Dynamics. *ACS Med. Chem. Lett.* **2012,** 3, 834-838.
34. Zhang, J. M.; Adrian, F. J.; Jahnke, W.; Cowan-Jacob, S. W.; Li, A. G.; Iacob, R. E.; Sim, T.; Powers, J.; Dierks, C.; Sun, F. X.; Guo, G. R.; Ding, Q.; Okram, B.; Choi, Y.; Wojciechowski, A.; Deng, X. M.; Liu, G. X.; Fendrich, G.; Strauss, A.; Vajpai, N.; Grzesiek, S.; Tuntland, T.; Liu, Y.; Bursulaya, B.; Azam, M.; Manley, P. W.; Engen, J. R.; Daley, G. Q.; Warmuth, M.; Gray, N. S. Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors. *Nature* **2010,** 463, 501-U116.
35. Adrian, F. J.; Ding, Q.; Sim, T. B.; Velentza, A.; Sloan, C.; Liu, Y.; Zhang, G. B.; Hur, W.; Ding, S.; Manley, P.; Mestan, J.; Fabbro, D.; Gray, N. S. Allosteric inhibitors of Bcr-abl-dependent cell proliferation. *Nat. Chem. Biol.* **2006,** 2, 95-102.
36. Liu, Q. S.; Sabnis, Y.; Zhao, Z.; Zhang, T. H.; Buhrlage, S. J.; Jones, L. H.; Gray, N. S. Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chem. Biol.* **2013,** 20, 146-159.
37. Garuti, L.; Roberti, M.; Bottegoni, G. Irreversible Protein Kinase Inhibitors. *Curr. Med. Chem.* **2011,** 18, 2981-2994.
38. Pan, Z. Y.; Scheerens, H.; Li, S. J.; Schultz, B. E.; Sprengeler, P. A.; Burrill, L. C.; Mendonca, R. V.; Sweeney, M. D.; Scott, K. C. K.; Grothaus, P. G.; Jeffery, D. A.; Spoerke, J. M.; Honigberg, L. A.; Young, P. R.; Dalrymple, S. A.; Palmer, J. T. Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. *Chemmedchem* **2007,** 2, 58-61.
39. Schlessinger, J. Receptor Tyrosine Kinases: Legacy of the First Two Decades. *Cold Spring Harb. Perspect. Biol.* **2014,** 6.
40. Lisabeth, E. M.; Falivelli, G.; Pasquale, E. B. Eph Receptor Signaling and Ephrins. *Cold Spring Harb. Perspect. Biol.* **2013,** 5.
41. Himanen, J. P.; Nikolov, D. B. Eph receptors and ephrins. *Int. J. Biochem. Cell Biol.* **2003,** 35, 130-134.
42. Pasquale, E. B. EPH receptor signalling casts a wide net on cell behaviour. *Nat. Rev. Mol. Cell Bio.* **2005,** 6, 462-475.
43. Gale, N. W.; Holland, S. J.; Valenzuela, D. M.; Flenniken, A.; Pan, L.; Ryan, T. E.; Henkemeyer, M.; Strebhardt, K.; Hirai, H.; Wilkinson, D. G.; Pawson, T.; Davis, S.; Yancopoulos, G. D. Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. *Neuron* **1996,** 17, 9-19.
44. Xia, G. B.; Kumar, S. R.; Masood, R.; Koss, M.; Templeman, C.; Quinn, D.; Zhu, S. T.; Reddy, R.; Krasnoperov, V.; Gill, P. S. Up-regulation of EphB4 in mesothelioma and its biological significance. *Clin. Cancer. Res.* **2005,** 11, 4305-4315.
45. Xia, G. B.; Kumar, S. R.; Masood, R.; Zhu, S. T.; Reddy, R.; Krasnoperov, V.; Quinn, D. I.; Henshall, S. M.; Sutherland, R. L.; Pinski, J. K.; Daneshmand, S.; Buscarini, M.; Stein, J. P.; Zhong, C.; Broek, D.; Roy-Burman, P.; Gill, P. S. EphB4 expression and biological significance in prostate cancer. *Cancer Res.* **2005,** 65, 4623-4632.
46. Berclaz, G.; Karamitopoulou, E.; Mazzucchelli, L.; Rohrbach, V.; Dreher, E.; Ziemiecki, A.; Andres, A. C. Activation of the receptor protein tyrosine kinase EphB4 in endometrial hyperplasia and endometrial carcinoma. *Ann. Oncol.* **2003,** 14, 220-226.

47. Huang, X.; Yamada, Y.; Kidoya, H.; Naito, H.; Nagahama, Y.; Kong, L.; Katoh, S. Y.; Li, W. F.; Ueno, M.; Takakura, N. EphB4 overexpression in B16 melanoma cells affects arterial-venous patterning in tumor angiogenesis. *Cancer Res.* **2007,** 67, 9800-9808.
48. Noren, N. K.; Lu, M.; Freeman, A. L.; Koolpe, M.; Pasquale, E. B. Interplay between EphB4 on tumor cells and vascular ephrin-B2 regulates tumor growth. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, 5583-5588.
49. Kumar, S. R.; Singh, J.; Xia, G.; Krasnoperov, V.; Hassanieh, L.; Ley, E. J.; Scehnet, J.; Kumar, N. G.; Hawes, D.; Press, M. F.; Weaver, F. A.; Gill, P. S. Receptor tyrosine kinase EphB4 is a survival factor in breast cancer. *Am. J. Pathol.* **2006,** 169, 279-93.
50. Batlle, E.; Bacani, J.; Begthel, H.; Jonkeer, S.; Gregorieff, A.; van de Born, M.; Malats, N.; Sancho, E.; Boon, E.; Pawson, T.; Gallinger, S.; Pals, S.; Clevers, H. EphB receptor activity suppresses colorectal cancer progression. *Nature* **2005,** 435, 1126-1130.
51. Noren, N. K.; Foos, G.; Hauser, C. A.; Pasquale, E. B. The EphB4 receptor suppresses breast cancer cell tumorigenicity through an Abl-Crk pathway. *Nat. Cell Biol.* **2006,** 8, 815-U53.
52. Dopeso, H.; Mateo-Lozano, S.; Mazzolini, R.; Rodrigues, P.; Lagares-Tena, L.; Ceron, J.; Romero, J.; Esteves, M.; Landolfi, S.; Hernandez-Losa, J.; Castano, J.; Wilson, A. J.; Cajal, S. R. Y.; Mariadason, J. M.; Schwartz, S.; Arango, D. The Receptor Tyrosine Kinase EPHB4 Has Tumor Suppressor Activities in Intestinal Tumorigenesis. *Cancer Res.* **2009,** 69, 7430-7438.
53. Kertesz, N.; Krasnoperov, V.; Reddy, R.; Leshanski, L.; Kumar, S. R.; Zozulya, S.; Gill, P. S. The soluble extracellular domain of EphB4 (sEphB4) antagonizes EphB4-EphrinB2 interaction, modulates angiogenesis, and inhibits tumor growth. *Blood* **2006,** 107, 2330-2338.
54. Martiny-Baron, G.; Korff, T.; Schaffner, F.; Esser, N.; Eggstein, S.; Marme, D.; Augustin, H. G. Inhibition of tumor growth and angiogenesis by soluble EphB4. *Neoplasia* **2004,** 6, 248-257.
55. Koolpe, M.; Burgess, R.; Dail, M.; Pasquale, E. B. EphB receptor-binding peptides identified by phage display enable design of an antagonist with ephrin-like affinity. *J. Biol. Chem.* **2005,** 280, 17301-17311.
56. Chrencik, J. E.; Brooun, A.; Recht, M. I.; Kraus, M. L.; Koolpe, M.; Kolatkar, A. R.; Bruce, R. H.; Martiny-Baron, G.; Widmer, H.; Pasquale, E. B.; Kuhn, P. Structure and thermodynamic characterization of the EphB4/ephrin-B2 antagonist peptide complex reveals the determinants for receptor specificity. *Structure* **2006,** 14, 321-330.
57. Miyazaki, Y.; Nakano, M.; Sato, H.; Truesdale, A. T.; Stuart, J. D.; Nartey, E. N.; Hightower, K. E.; Kane-Carson, L. Design and effective synthesis of novel templates, 3,7-diphenyl-4amino-thieno and furo-[3,2-c]pyridines as protein kinase inhibitors and in vitro evaluation targeting angiogenetic kinases. *Bioorg. Med. Chem. Lett.* **2007,** 17, 250-254.
58. Bardelle, C.; Cross, D.; Davenport, S.; Kettle, J. G.; Ko, E. J.; Leach, A. G.; Mortlock, A.; Read, J.; Roberts, N. J.; Robins, P.; Williams, E. J. Inhibitors of the tyrosine kinase EphB4. Part 1: Structure-based design and optimization of a series of 2,4-bis-anilinopyrimidines. *Bioorg. Med. Chem. Lett.* **2008,** 18, 2776-2780.

59. Bardelle, C.; Coleman, T.; Cross, D.; Davenport, S.; Kettle, J. G.; Ko, E. J.; Leach, A. G.; Mortlock, A.; Read, J.; Roberts, N. J.; Robins, P.; Williams, E. J. Inhibitors of the tyrosine kinase EphB4. Part 2: Structure-based discovery and optimisation of 3,5-bis substituted anilinopyrimidines. *Bioorg. Med. Chem. Lett.* **2008,** 18, 5717-5721.

60. Bardelle, C.; Barlaam, B.; Brooks, N.; Coleman, T.; Cross, D.; Ducray, R.; Green, I.; Lambert-van der Brempt, C.; Olivier, A.; Read, J. Inhibitors of the tyrosine kinase EphB4. Part 3: Identification of non-benzodioxole-based kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2010,** 20, 6242-6245.

61. Barlaam, B.; Ducray, R.; Lambert-van der Brempt, C.; Ple, P.; Bardelle, C.; Brooks, N.; Coleman, T.; Cross, D.; Kettle, J. G.; Read, J. Inhibitors of the tyrosine kinase EphB4. Part 4: Discovery and optimization of a benzylic alcohol series. *Bioorg. Med. Chem. Lett.* **2011,** 21, 2207-2211.

62. Mitchell, S. A.; Danca, M. D.; Blomgren, P. A.; Darrow, J. W.; Currie, K. S.; Kropf, J. E.; Lee, S. H.; Gallion, S. L.; Xiong, J. M.; Pippin, D. A.; DeSimone, R. W.; Brittelli, D. R.; Eustice, D. C.; Bourret, A.; Hill-Drzewi, M.; Maciejewski, P. M.; Elkin, L. L. Imidazo[1,2-a]pyrazine diaryl ureas: Inhibitors of the receptor tyrosine kinase EphB4. *Bioorg. Med. Chem. Lett.* **2009,** 19, 6991-6995.

63. Qiao, L. X.; Choi, S.; Case, A.; Gainer, T. G.; Seyb, K.; Glicksman, M. A.; Lo, D. C.; Stein, R. L.; Cuny, G. D. Structure-activity relationship study of EphB3 receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2009,** 19, 6122-6126.

64. Martiny-Baron, G.; Holzer, P.; Billy, E.; Schnell, C.; Brueggen, J.; Ferretti, M.; Schmiedeberg, N.; Wood, J. M.; Furet, P.; Imbach, P. The small molecule specific EphB4 kinase inhibitor NVP-BHG712 inhibits VEGF driven angiogenesis. *Angiogenesis* **2010,** 13, 259-267.

65. Wang, Y.; Shakespeare, W. C.; Huang, W. S.; Sundaramoorthi, R.; Lentini, S.; Das, S.; Liu, S.; Banda, G.; Wen, D.; Zhu, X.; Xu, Q.; Keats, J.; Wang, F.; Wardwell, S.; Ning, Y.; Snodgrass, J. T.; Broudy, M. I.; Russian, K.; Dalgarno, D.; Clackson, T.; Sawyer, T. K. Novel N9-arenethenyl purines as potent dual Src/Abl tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2008,** 18, 4907-12.

66. Stroylov, V. S.; Rakitina, T. V.; Novikov, F. N.; Stroganov, O. V.; Chilov, G. G.; Lipkin, A. V. Novel fragment-like inhibitors of EphA2 obtained by experimental screening and modelling. *Mendeleev Commun.* **2010,** 20, 263-265.

67. Chang, Q.; Jorgensen, C.; Pawson, T.; Hedley, D. W. Effects of dasatinib on EphA2 receptor tyrosine kinase activity and downstream signalling in pancreatic cancer. *Br. J. Cancer* **2008,** 99, 1074-1082.

68. Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008,** 26, 127-32.

69. Type XIII RTKs: Ephrin receptor family: EPH receptor B4. Last modified on 26/03/2013. Accessed on 12/10/2014. IUPHAR database (IUPHAR-DB), http://www.iuphar-db.org/DATABASE/ObjectDisplayForward?objectId=1833.

70. Melnick, J. S.; Janes, J.; Kim, S.; Chang, J. Y.; Sipes, D. G.; Gunderson, D.; Jarnes, L.; Matzen, J. T.; Garcia, M. E.; Hood, T. L.; Beigi, R.; Xia, G.; Harig, R. A.; Asatryan, H.; Yan, S. F.; Zhou, Y. Y.; Gu, X. J.; Saadat, A.; Zhou, V.; King, F. J.; Shaw, C. M.; Su, A. I.; Downs, R.; Gray, N. S.; Schultz, P. G.; Warmuth, M.; Caldwell, J. S. An efficient rapid system for profiling the cellular activities of molecular libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** 103, 3153-3158.

71. Gendreau, S. B.; Ventura, R.; Keast, P.; Laird, A. D.; Yakes, F. M.; Zhang, W. T.; Bentzien, F.; Cancilla, B.; Lutman, J.; Chu, F.; Jackman, L.; Shi, Y. C.; Yu, P. W.; Wang, J.; Aftab, D. T.; Jaeger, C. T.; Meyer, S. M.; De Costa, A.; Engell, K.; Chen, J.; Martini, J. F.; Joly, A. H. Inhibition of the T790M gatekeeper mutant of the epidermal growth factor receptor by EXEL-7647. *Clin. Cancer. Res.* **2007,** 13, 3713-3723.

72. Caligiuri, M.; Molz, L.; Liu, Q.; Kaplan, F.; Xu, J. P.; Majeti, J. Z.; Ramos-Kelsey, R.; Murthi, K.; Lievens, S.; Tavernier, J.; Kley, N. MASPIT: Three-hybrid trap for quantitative proteome fingerprinting of small molecule-protein interactions in mammalian cells. *Chem. Biol.* **2006,** 13, 711-722.

73. Burkholder, T. P.; Clayton, J. R.; Rempala, M. E.; Henry, J. R.; Knobeloch, J. M.; Mendel, D.; McLean, J. A.; Hao, Y.; Barda, D. A.; Considine, E. L.; Uhlik, M. T.; Chen, Y. F.; Ma, L. D.; Bloem, L. J.; Akunda, J. K.; McCann, D. J.; Sanchez-Felix, M.; Clawson, D. K.; Lahn, M. M.; Starling, J. J. Discovery of LY2457546: a multi-targeted anti-angiogenic kinase inhibitor with a novel spectrum of activity and exquisite potency in the acute myelogenous leukemia-Flt-3-internal tandem duplication mutant human tumor xenograft model. *Invest. New Drugs* **2012,** 30, 936-949.

74. Bain, J.; Plater, L.; Elliott, M.; Shpiro, N.; Hastie, C. J.; Mclauchlan, H.; Klevernic, I.; Arthur, J. S. C.; Alessi, D. R.; Cohen, P. The selectivity of protein kinase inhibitors: a further update. *Biochem. J.* **2007,** 408, 297-315.

75. Sturz, A.; Bader, B.; Thierauch, K. H.; Glienke, J. EphB4 signaling is capable of mediating ephrinB2-induced inhibition of cell migration. *Biochem. Biophys. Res. Commun.* **2004,** 313, 80-88.

76. Kolb, P.; Kipouros, C. B.; Huang, D. Z.; Caflisch, A. Structure-based tailoring of compound libraries for high-throughput screening: Discovery of novel EphB4 kinase inhibitors. *Proteins-Structure Function and Bioinformatics* **2008,** 73, 11-18.

77. Zhou, T.; Caflisch, A. High-Throughput Virtual Screening Using Quantum Mechanical Probes: Discovery of Selective Kinase Inhibitors. *Chemmedchem* **2010,** 5, 1007-1014.

78. Dong, J.; Zhao, H.; Zhou, T.; Spiliotopoulos, D.; Rajendran, C.; Li, X.; Huang, D.; Caflisch, A. Structural Analysis of the Binding of Type I, I1/2, and II Inhibitors to Eph Tyrosine Kinases. *ACS Med. Chem. Lett.* **2014,** Article ASAP, DOI: 10.1021/ml500355x.

79. Zhao, H. T.; Huang, D. Z.; Caflisch, A. Discovery of Tyrosine Kinase Inhibitors by Docking into an Inactive Kinase Conformation Generated by Molecular Dynamics. *Chemmedchem* **2012,** 7, 1983-1990.

80. Zhao, H.; Huang, D. Hydrogen bonding penalty upon ligand binding. *PLoS One* **2011,** 6, e19923.

81. Boyd, A. W.; Bartlett, P. F.; Lackmann, M. Therapeutic targeting of EPH receptors and their ligands. *Nat. Rev. Drug Discov.* **2014,** 13, 39-62.

82. Bold, G.; Altmann, K. H.; Frei, J.; Lang, M.; Manley, P. W.; Traxler, P.; Wietfeld, B.; Bruggen, J.; Buchdunger, E.; Cozens, R.; Ferrari, S.; Furet, P.; Hofmann, F.; Martiny-Baron, G.; Mestan, J.; Rosel, J.; Sills, M.; Stover, D.; Acemoglu, F.; Boss, E.; Emmenegger, R.; Lasser, L.; Masso, E.; Roth, R.; Schlachter, C.; Vetterli, W. New anilinophthalazines as potent and orally well absorbed inhibitors of the VEGF receptor tyrosine kinases useful as antagonists of tumor-driven angiogenesis. *J. Med. Chem.* **2000,** 43, 2310-23.

83. Smith, R. A.; Barbosa, J.; Blum, C. L.; Bobko, M. A.; Caringal, Y. V.; Dally, R.; Johnson, J. S.; Katz, M. E.; Kennure, N.; Kingery-Wood, J.; Lee, W.; Lowinger, T. B.; Lyons, J.; Marsh, V.; Rogers, D. H.; Swartz, S.; Walling, T.; Wild, H. Discovery of heterocyclic ureas as a new class of raf kinase inhibitors: identification of a second generation lead by a combinatorial chemistry approach. *Bioorg. Med. Chem. Lett.* **2001,** 11, 2775-8.

84. Khire, U. R.; Bankston, D.; Barbosa, J.; Brittelli, D. R.; Caringal, Y.; Carlson, R.; Dumas, J.; Gane, T.; Heald, S. L.; Hibner, B.; Johnson, J. S.; Katz, M. E.; Kennure, N.; Kingery-Wood, J.; Lee, W.; Liu, X. G.; Lowinger, T. B.; McAlexander, I.; Monahan, M. K.; Natero, R.; Renick, J.; Riedl, B.; Rong, H.; Sibley, R. N.; Smith, R. A.; Wolanin, D. Omega-carboxypyridyl substituted ureas as Raf kinase inhibitors: SAR of the amide substituent. *Bioorg. Med. Chem. Lett.* **2004,** 14, 783-6.

85. Albaugh, P.; Fan, Y.; Mi, Y.; Sun, F. X.; Adrian, F.; Li, N. X.; Jia, Y.; Sarkisova, Y.; Kreusch, A.; Hood, T.; Lu, M.; Liu, G. X.; Huang, S. L.; Liu, Z. S.; Loren, J.; Tuntland, T.; Karanewsky, D. S.; Seidel, H. M.; Molteni, V. Discovery of GNF-5837, a Selective TRK Inhibitor with Efficacy in Rodent Cancer Tumor Models. *ACS Med. Chem. Lett.* **2012,** 3, 140-145.

86. Liu, C.; Lin, J.; Wrobleski, S. T.; Lin, S.; Hynes, J.; Wu, H.; Dyckman, A. J.; Li, T.; Wityak, J.; Gillooly, K. M.; Pitt, S.; Shen, D. R.; Zhang, R. F.; McIntyre, K. W.; Salter-Cid, L.; Shuster, D. J.; Zhang, H.; Marathe, P. H.; Doweyko, A. M.; Sack, J. S.; Kiefer, S. E.; Kish, K. F.; Newitt, J. A.; McKinnon, M.; Dodd, J. H.; Barrish, J. C.; Schieven, G. L.; Leftheris, K. Discovery of 4-(5-(cyclopropylcarbamoyl)-2-methylphenylamino)-5-methyl-N-propylpyrrolo[1,2-f][ 1,2,4]triazine-6-carboxamide (BMS-582949), a clinical p38alpha MAP kinase inhibitor for the treatment of inflammatory diseases. *J. Med. Chem.* **2010,** 53, 6629-39.

87. Hynes, J., Jr.; Wu, H.; Pitt, S.; Shen, D. R.; Zhang, R.; Schieven, G. L.; Gillooly, K. M.; Shuster, D. J.; Taylor, T. L.; Yang, X.; McIntyre, K. W.; McKinnon, M.; Zhang, H.; Marathe, P. H.; Doweyko, A. M.; Kish, K.; Kiefer, S. E.; Sack, J. S.; Newitt, J. A.; Barrish, J. C.; Dodd, J.; Leftheris, K. The discovery of (R)-2-(sec-butylamino)-N-(2-methyl-5-(methylcarbamoyl)phenyl) thiazole-5-carboxamide (BMS-640994)-A potent and efficacious p38alpha MAP kinase inhibitor. *Bioorg. Med. Chem. Lett.* **2008,** 18, 1762-7.

88. Wrobleski, S. T.; Lin, S.; Dhar, T. G.; Dyckman, A. J.; Li, T.; Pitt, S.; Zhang, R.; Fan, Y.; Doweyko, A. M.; Tokarski, J. S.; Kish, K. F.; Kiefer, S. E.; Sack, J. S.; Newitt, J. A.; Witmer, M. R.; McKinnon, M.; Barrish, J. C.; Dodd, J. H.; Schieven, G. L.; Leftheris, K. The identification of novel p38alpha isoform selective kinase inhibitors having an unprecedented p38alpha binding mode. *Bioorg. Med. Chem. Lett.* **2013,** 23, 4120-6.

89. Weisberg, E.; Manley, P. W.; Breitenstein, W.; Bruggen, J.; Cowan-Jacob, S. W.; Ray, A.; Huntly, B.; Fabbro, D.; Fendrich, G.; Hall-Meyers, E.; Kung, A. L.; Mestan, J.; Daley, G. Q.; Callahan, L.; Catley, L.; Cavazza, C.; Azam, M.; Neuberg, D.; Wright, R. D.; Gilliland, D. G.; Griffin, J. D. Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. *Cancer Cell* **2005,** 7, 129-41.

90. Ishikawa, M.; Hashimoto, Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. *J. Med. Chem.* **2011,** 54, 1539-54.

91. Pratt, E. F.; Kereszte.Jc. Syntheses of Indolizino- and Dihydroindolizinoquinoxalines. *J. Org. Chem.* **1967,** 32, 49-&.

92. Obafemi, C. A.; Pfleiderer, W. Synthesis and some reactions of 3-chloro-2-(cyanomethylene)-1,2-dihydroquinoxalines. *Molecules* **2004,** 9, 223-31.

93. Otomasu, H.; Ohmiya, S.; Sekuguch.T; Takahash.H. Synthesis of Condensed Quinoxalines .2. A New Synthesis of Pyrrolo- 2,3-B!Quinoxalines. *Chem. Pharm. Bull. (Tokyo)* **1970,** 18, 2065-&.

94. Choi, H. G.; Ren, P.; Adrian, F.; Sun, F.; Lee, H. S.; Wang, X.; Ding, Q.; Zhang, G.; Xie, Y.; Zhang, J.; Liu, Y.; Tuntland, T.; Warmuth, M.; Manley, P. W.; Mestan, J.; Gray, N. S.; Sim, T. A type-II kinase inhibitor capable of inhibiting the T315I "gatekeeper" mutant of Bcr-Abl. *J. Med. Chem.* **2010,** 53, 5439-48.

95. Huang, W. S.; Shakespeare, W. C. An efficient synthesis of nilotinib (AMN107). *Synthesis-Stuttgart* **2007**, 2121-2124.

96. Jung, M. E.; Ouk, S.; Yoo, D.; Sawyers, C. L.; Chen, C.; Tran, C.; Wongvipat, J. Structure-activity relationship for thiohydantoin androgen receptor antagonists for castration-resistant prostate cancer (CRPC). *J. Med. Chem.* **2010,** 53, 2779-96.

97. Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007,** 2, 2212-21.

98. Bamborough, P.; Angell, R. M.; Bhamra, I.; Brown, D.; Bull, J.; Christopher, J. A.; Cooper, A. W.; Fazal, L. H.; Giordano, I.; Hind, L.; Patel, V. K.; Ranshaw, L. E.; Sims, M. J.; Skone, P. A.; Smith, K. J.; Vickerstaff, E.; Washington, M. N-4-Pyrimidinyl-1H-indazol-4-amine inhibitors of Lck: indazoles as phenol isosteres with improved pharmacokinetics. *Bioorg. Med. Chem. Lett.* **2007,** 17, 4363-8.

99. Graham Robinett, R.; Freemerman, A. J.; Skinner, M. A.; Shewchuk, L.; Lackey, K. The discovery of substituted 4-(3-hydroxyanilino)-quinolines as potent RET kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2007,** 17, 5886-93.

100. Hynes, J., Jr.; Dyckman, A. J.; Lin, S.; Wrobleski, S. T.; Wu, H.; Gillooly, K. M.; Kanner, S. B.; Lonial, H.; Loo, D.; McIntyre, K. W.; Pitt, S.; Shen, D. R.; Shuster, D. J.; Yang, X.; Zhang, R.; Behnia, K.; Zhang, H.; Marathe, P. H.; Doweyko, A. M.; Tokarski, J. S.; Sack, J. S.; Pokross, M.; Kiefer, S. E.; Newitt, J. A.; Barrish, J. C.; Dodd, J.; Schieven, G. L.; Leftheris, K. Design, synthesis, and anti-inflammatory properties of orally active 4-(phenylamino)-pyrrolo[2,1-f][1,2,4]triazine p38alpha mitogen-activated protein kinase inhibitors. *J. Med. Chem.* **2008,** 51, 4-16.

101. Deak, H. L.; Newcomb, J. R.; Nunes, J. J.; Boucher, C.; Cheng, A. C.; DiMauro, E. F.; Epstein, L. F.; Gallant, P.; Hodous, B. L.; Huang, X.; Lee, J. H.; Patel, V. F.; Schneider, S.; Turci, S. M.; Zhu, X. N-(3-(phenylcarbamoyl)arylpyrimidine)-5-

carboxamides as potent and selective inhibitors of Lck: structure, synthesis and SAR. *Bioorg. Med. Chem. Lett.* **2008,** 18, 1172-6.
102. Kufareva, I.; Abagyan, R. Type-II kinase inhibitor docking, screening, and profiling using modified structures of active kinase states. *J. Med. Chem.* **2008,** 51, 7921-32.
103. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* **2006,** 5, 730-9.
104. Paul W. Manley, S. W. C.-J., Gabriele Fendrich, Wolfgang Jahnke and Doriano Fabbro. Nilotinib, in Comparison to Both Dasatinib and Imatinib, Possesses a Greatly Prolonged Residence Time When Bound to the BCR-ABL Kinase SH1 Domain. In *53rd ASH Annual Meeting and Exposition*, San Diego, USA, 2011.
105. Kitagawa, D.; Gouda, M.; Kirii, Y. Quick evaluation of kinase inhibitors by surface plasmon resonance using single-site specifically biotinylated kinases. *J. Biomol. Screen.* **2014,** 19, 453-61.
106. Wood, E. R.; Truesdale, A. T.; McDonald, O. B.; Yuan, D.; Hassell, A.; Dickerson, S. H.; Ellis, B.; Pennisi, C.; Horne, E.; Lackey, K.; Alligood, K. J.; Rusnak, D. W.; Gilmer, T. M.; Shewchuk, L. A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res.* **2004,** 64, 6652-9.
107. Fabian, M. A.; Biggs, W. H., 3rd; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lelias, J. M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **2005,** 23, 329-36.
108. Stephenson, S. A.; Slomka, S.; Douglas, E. L.; Hewett, P. J.; Hardingham, J. E. Receptor protein tyrosine kinase EphB4 is up-regulated in colon cancer. *BMC Mol. Biol.* **2001,** 2, 15.
109. Castellano, G.; Reid, J. F.; Alberti, P.; Carcangiu, M. L.; Tomassetti, A.; Canevari, S. New potential ligand-receptor signaling loops in ovarian cancer identified in multiple gene expression studies. *Cancer Res.* **2006,** 66, 10709-19.
110. Unzue, A.; Dong, J.; Lafleur, K.; Zhao, H. T.; Frugier, E.; Caflisch, A.; Nevado, C. Pyrrolo[3,2-b]quinoxaline Derivatives as Types I-1/2 and II Eph Tyrosine Kinase Inhibitors: Structure-Based Design, Synthesis, and in Vivo Validation. *J. Med. Chem.* **2014,** 57, 6834-6844.
111. Adams, R. H. Vascular patterning by Eph receptor tyrosine kinases and ephrins. *Semin. Cell Dev. Biol.* **2002,** 13, 55-60.
112. Korff, T.; Augustin, H. G. Tensional forces in fibrillar extracellular matrices control directional capillary sprouting. *J. Cell Sci.* **1999,** 112 ( Pt 19), 3249-58.
113. Pratt, E. F.; Kereszte, J. C. Syntheses of Indolizino- and Dihydroindolizinoquinoxalines. *J. Org. Chem.* **1967,** 32, 49-53.
114. Choi, H. G.; Ren, P. D.; Adrian, F.; Sun, F. X.; Lee, H. S.; Wang, X.; Ding, Q. A.; Zhang, G. B.; Xie, Y. P.; Zhang, J. M.; Liu, Y.; Tuntland, T.; Warmuth, M.; Manley, P. W.; Mestan, J.; Gray, N. S.; Sim, T. A Type-II Kinase Inhibitor Capable of Inhibiting the T315I "Gatekeeper" Mutant of Bcr-Abl. *J. Med. Chem.* **2010,** 53, 5439-5448.

115. Jung, M. E.; Ouk, S.; Yoo, D.; Sawyers, C. L.; Chen, C.; Tran, C.; Wongvipat, J. Structure-Activity Relationship for Thiohydantoin Androgen Receptor Antagonists for Castration-Resistant Prostate Cancer (CRPC). *J. Med. Chem.* **2010,** 53, 2779-2796.
116. Aquila, B.; Ezhuthachan, J.; Dakin, L.; Lyne, P.; Lee, S.; Pontz, T.; Zheng, X. Quinazoline derivatives and their use as B-Raf inhibitors. 2009.
117. Betebenner, D. A.; Degoey, D. A.; Marning, C. J.; Krueger, A. C.; Iwasaki, N.; Rockway, T. W.; Cooper, C. S.; Anderson, D. D.; Donner, P. L.; Green, B. E.; Kempf, D. J.; Liu, D.; McDaniel, K. F.; Madigan, D. L.; Motter, C. E.; Pratt, J. K.; Shanley, J. P.; Tufano, M. D.; Wagner, R.; Zhang, R.; Molla, A.; Mo, H.; Pilot-Matias, T. J.; Masse, S. V.; Carrick, R. J.; He, W.; Lu, L.; Gampovnik, D. J. Anti-Viral compounds. 2007.
118. Deak, H. L.; Newcomb, J. R.; Nunes, J. J.; Boucher, C.; Cheng, A. C.; DiMauro, E. F.; Epstein, L. F.; Gallant, P.; Hodous, B. L.; Huang, X.; Lee, J. H.; Patel, V. F.; Schneider, S.; Turci, S. M.; Zhu, X. T. N-(3-(Phenylcarbamoyl)arylpyrimidine)-5-carboxamides as potent and selective inhibitors of Lck: Structure, synthesis and SAR. *Bioorg. Med. Chem. Lett.* **2008,** 18, 1172-1176.
119. Oguro, Y.; Miyamoto, N.; Okada, K.; Takagi, T.; Iwata, H.; Awazu, Y.; Miki, H.; Hori, A.; Kamiyama, K.; Imamura, S. Design, synthesis, and evaluation of 5-methyl-4-phenoxy-5H-pyrrolo[3,2-d]pyrimidine derivatives: Novel VEGFR2 kinase inhibitors binding to inactive kinase conformation. *Biorg. Med. Chem.* **2010,** 18, 7260-7273.
120. Choi, Y.; Syeda, F.; Walker, J. R.; Finerty, P. J.; Cuerrier, D.; Wojciechowski, A.; Liu, Q. S.; Dhe-Paganon, S.; Gray, N. S. Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2009,** 19, 4467-4470.
121. Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J. Appl. Crystallogr.* **1993,** 26, 795-800.
122. Bailey, S. The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallogr D* **1994,** 50, 760-763.
123. Mccoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007,** 40, 658-674.
124. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D* **2002,** 58, 1948-1954.
125. Korff, T.; Augustin, H. G. Tensional forces in fibrillar extracellular matrices control directional capillary sprouting. *J. Cell Sci.* **1999,** 112, 3249-3258.
126. Korff, T.; Augustin, H. G. Integration of endothelial cells in multicellular spheroids prevents apoptosis and induces differentiation. *J. Cell Biol.* **1998,** 143, 1341-1352.

# 3

# BROMODOMAIN LIGANDS

## 3.1 Introduction

Bromodomain proteins constitute an emerging topic in the field of drug discovery due, in a large extent, to their involvement in the regulation of many genes.[1] In the next sections, epigenetic targets are introduced, with a special focus on bromodomains and their ligands.

### 3.1.1 Epigenetics

The central dogma of molecular biology, enunciated by Francis Crick in 1956 and re-stated in 1970, describes the flow of genetic information in living organisms: how genes in DNA direct the synthesis of RNA and proteins, as well as how DNA is replicated (Figure 1).[2] Transcription is the first step of gene expression, where a DNA segment is converted into mRNA by RNA polymerases. The mRNA is then translated in the ribosomes into proteins, and as a consequence, the information changes from being stored as nucleotide base pair sequences to aminoacid sequences.

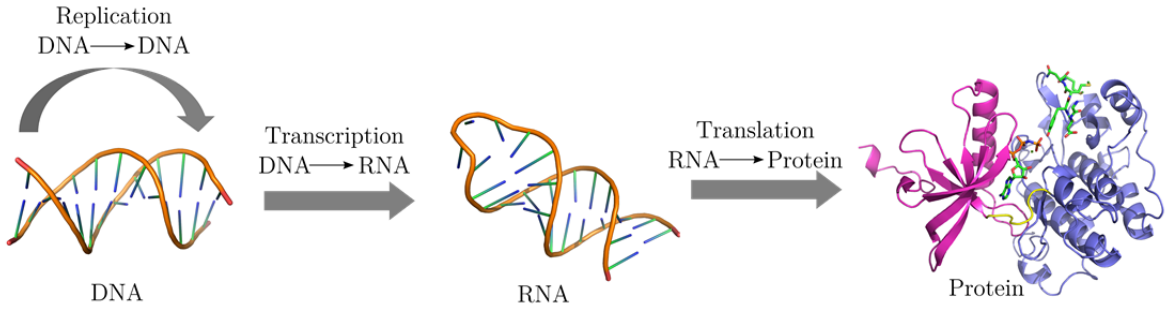



**Figure 1.** Schematic representation of the dogma of molecular biology.

Epigenetics, from the greek *Epi*, which means *above*, are mechanisms that regulate gene transcription that are not encoded by the DNA sequence, and has been defined as ‘‘regulation

beyond our genes''.[3] Not only our genetic predispositions play a key role on the development of certain diseases, but so do the more dynamic epigenetic processes during DNA transcription. Modifying our genome remains a highly challenging endeavor, but there is an emerging possibility to control and modify epigenetic motifs that might be key for the development of certain diseases. More importantly, it has been shown that cancer initiation and progression occurs not only because of genetic modifications, but also due to epigenetic alterations.[4-5]

The genetic material is compactly stored within the nucleus of the cell in chromosomes, which are condensed in chromatin fibers. Within chromatin, the double helix DNA is wrapped around nucleosomes (Figure 2). Nucleosomes are octamer proteins composed of four histone proteins, whose flexible N-terminus is called ''histone tail''. The surface of histones is mainly occupied by aminoacids containing basic side chains, such as lysines and asparagines, which are positively charged at physiological pH. As DNA is negatively charged, it can easily bind to the positively charged histone surface establishing salt bridges, hydrogen bonds and hydrophobic interactions. As a consequence, DNA-based processes such as replication, transcription and repair are highly dependent on chromatin's structure (referred to as chromatin remodeling).

''Histone tails'' are subjected to a wide range of post transcriptional modifications (PTMs), which affect chromatin architecture (by altering histone-histone and histone-DNA interactions), gene expression, cell fate determination, differentiation and disease onset and/or maintenance.[6-7] Acetylation, mono/multi methylation and phosphorylation of aminoacid residues in the ''histone tails'' are post-translational modifications that are likely to occur and are enclosed in the so called epigenetic ''histone code'' (defined as the ensemble of PTMs occurring in histone proteins).[1, 8-9] Such modifications are a result of a concerted, sequence and context dependent action of different protein interaction modules that are classified in three major groups (Figure 2): ''writers'', ''erasers'' and ''readers''.

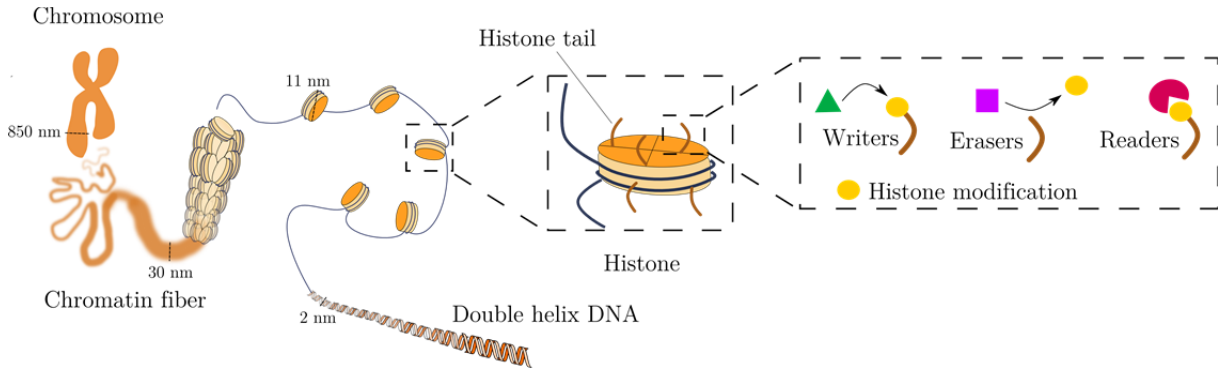



**Figure 2.** Chromatin is a complex composed of DNA, RNA and proteins that is located in the nucleus of the cell. It compacts DNA into a smaller volume to fit in the cell, allows cell mitosis, prevents DNA damage and controls gene expression and DNA replication. In chromatin the DNA is wrapped around histones, whose tails are subjected to multiple PTMs by ''writers'', ''erasers'' and ''readers''. The PTMs include lysine acetylation, which will be explained in detail in the next sub-section.[10]

The message that is brought by epigenetic ‘‘writers’’ such as HATs (histone acetyltransferases), HMTs (histone mehtyltransferases), kinases (see chapter 2 for more information) and ubiquitinases is deciphered by epigenetic ‘‘readers’’ such as bromodomains, chromodomains or tudors among others.[11-12] The opposite effect, the removal of the message, is carried out by the “erasers” (e.g. HDACs, histone deacetylases). In this chapter special attention is given to epigenetic ‘‘readers’’, specifically to bromodomains.

### I. Protein lysine acetylation

Protein lysine acetylation (KAc) is an important PTM within the ‘‘histone code’’[1, 8-9] that takes place on over 1750 proteins.[13] It occurs when an acetyl group of acetyl-CoA is transferred to the ε amino group of a lysine residue by a histone acetyltransferase (HAT) (Figure 3). Acetylation causes a neutralization of the charge in lysine residues that leads to a change in the electrostatic interaction between them and the negatively charged DNA.[3] By neutralizing the charge of the ‘‘histone tails’’, chromatin becomes more flexible, making DNA more accessible for transcription (Figure 1). Deregulation in the acetylation levels can result in changes in gene expression such as the inactivation of tumor suppressor functions and the activation of oncogenes, leading to cancer, among other diseases. As a result, lysine acetylation is nowadays considered an important PTM that is able to modulate several cellular proteins and its importance has been compared to protein phosphorylation performed by tyrosine kinases (see Chapter 2 for more information).

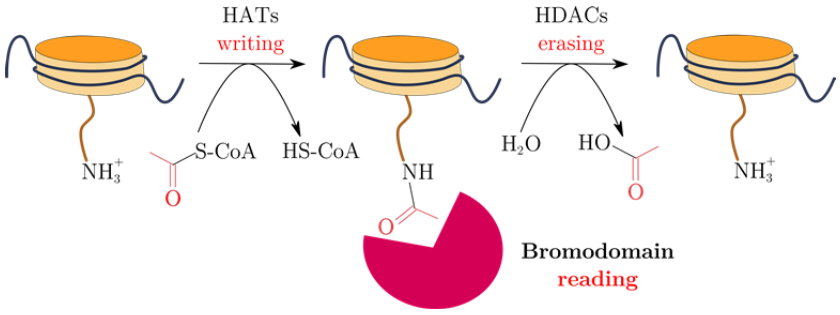



**Figure 3.** Histone acetyltransferase (HAT) enzymes (“writers”) introduce acetyl groups in lysines (left), which are “read” by bromodomains (middle). The acetylated residues are removed by histone deacetylases (HDACs), that act as “erasers” (right).[14]

The FDA (U. S. Food and Drug Administration) approved two drugs for the treatment of T-cell lymphoma called vorinostat[15] and romidepsin,[16] which are histone deacetylase (HDAC) inhibitors, thus demonstrating the potential to treat diseases by modifying acetylation pathways. As a consequence, there has been an increased interest in the search of small molecules that are able to interact with acetyl lysine ‘‘erasers’’, ‘‘writers’’ and recently also with ‘‘readers’’ (including bromodomains).[3]

Bromodomains are potential therapeutic targets due to their ability to modify cellular lysine acetylation[17-22] and have been associated with several diseases such as diabetes, obesity, oncology, inflammation, viral infection and metabolic and cardiovascular disorders.[23-26] The interest in bromodomain biology and the efforts towards the development of bromodomain ligands has dramatically increased over the last 3-4 years.[14, 27-28] However, bromodomain related research has just started, and the underlying biological processes and exact therapeutic pathways related to these proteins remain unknown.

### 3.1.2 Bromodomains: structure and function

Bromodomains are named after the Drosophila gene *brahma*, where they were first identified.[29] They are protein interaction modules that are part of large protein architectures and function as epigenetic ''readers'' (Figures 2, 3). Bromodomains specifically recognize the ε-N-acetylated lysine residues (KAc group) present in proteins (especially in histones)[30-31] through a conserved interaction mode, altering the process of chromatin remodelling.[32] However, bromodomains are only able to bind to acetylated lysines with moderate affinity presenting $K_d$ values in the range of tens to hundreds of micromolar.[33-34] There are 61 different bromodomains recognized within the human genome spread across 46 multidomain proteins (with up to six bromodomains per protein)[35-36] and they are phylogenetically divided into eight distinct families as shown in Figure 4.[37]

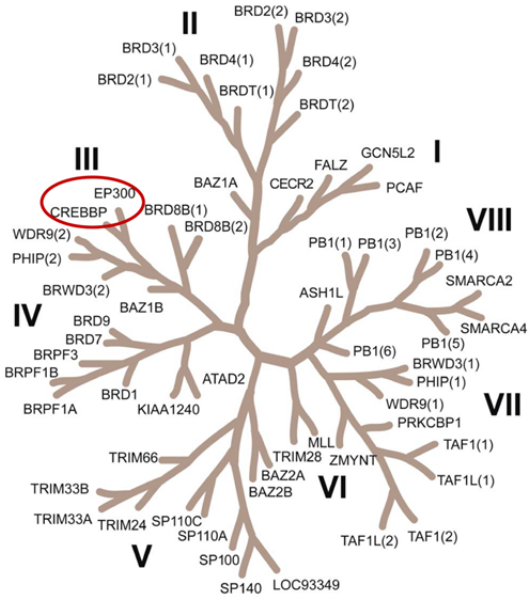



**Figure 4.** Phylogenetic tree of the human bromodomain family.[38] Bromodomains are phylogenetically divided into eight families (I-VIII). CREBBP and EP300, proteins of interest in this work, are highlighted in red. Family II, also known as BET family, is the most studied bromodomain family up to date.

Bromodomains are part of multidomain proteins that present specific functions. According to the function of the multidomain protein, bromodomains can be further classified into

different subgroups: histone acetyltransferases (known as HATs, such as CREBBP, GCN5 and PCAF, which function as transcription coactivators), histone methyltransferases (ASH1L, MLL), ATP-dependent chromatin-remodeling complexes (Swi2/Snf2) and transcriptional regulators carrying tandem bromodomains and an extra terminal domain (the BET family of bromodomains, family II, Figure 4).

Bromodomains consist of approximately 110 amino acids and despite their large sequence diversity, they share a common tertiary structure that is described as a left-handed bundle of four helices (αZ, αA, αB and αC), which are linked by two loops (ZA and BC loops).[39] The hydrophobic KAc binding site is conserved in all bromodomains and is formed by the ZA and BC loops at one end of the helical bundle (Figure 5, A, B).[31] It anchores the carbonyl oxygen of the acetyl group of the KAc by forming two hydrogen bonds; one to the $NH_2$ of the conserved asparagine (Asn1168 in the CREBBP bromodomain, Figure 5, C, hydrogen bonds shown as dashed lines) and a second one to the side chain of a conserved tyrosine residue (Tyr1125, Figure 5, C, hydrogen bonds shown as dashed lines)[31, 40] through one structured water molecule (Figure 5, B and C, shown as a spheres and ellipses respectively ).

The so called ZA channel, as well as the LPF shelf (indicated in Figure 5, B, C) have proven to be of high importance during the design of bromodomain ligands.[41-42] Both regions are located within the ZA loop, and the LPF shelf, in particular, is composed of three aminoacids, leucine (L), proline (P) and phenylalanine (F), which vary among different bromodomains (e.g. it corresponds to the WPF shelf, tryptophan, proline and phenylalanine, in the case of the BET bromodomain family).

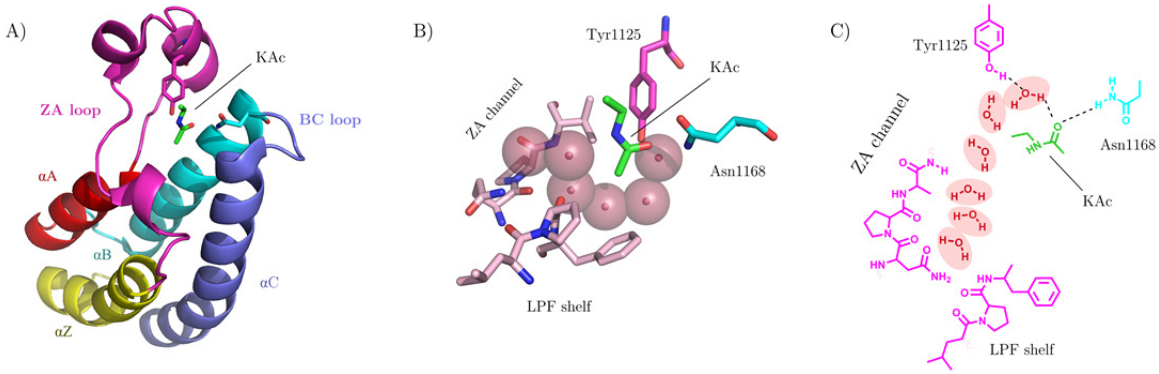



**Figure 5.** A) Ribbon representation of the CREBBP bromodomain composed of the four canonical helices (αZ, αA, αB and αC) and two loop regions (ZA and BC loops) in complex with KAc (PDB code 3P1C). B) Typical binding mode of KAc to bromodomains forming interactions, in the case of CREBBP, with Asn1168 (in light blue) and Tyr1125 (in magenta, via a structure water molecule, shown as spheres). The residues belonging to the ZA channel and LPF shelf, two important regions of the binding site, are indicated. C) 2D representation of the binding mode of the acetylated lysine KAc (in green). Water molecules are shown as ellipses and the hydrogen bonds are indicated as dashed lines. The residues belonging to the ZA channel and LPF shelf, two important regions of the binding site, are indicated.

The acetyl group of KAc interacts with a water molecule network composed of four to six water molecules located at the bottom of the binding site that provides stability to the KAc binding site (shown as spheres and ellipses, Figure 5, B and C respectively). The asparagine (Asn1168) and tyrosine residue (Tyr1125) that are essential for the function of the bromodomain also take part in the hydrogen-bonding network with the crystallographically ordered water molecules.

According to Molecular Dynamic simulations (MD) performed on several bromodomains[17, 43] the four protein helices are rather rigid, whereas the ZA and BC loop regions that contribute to KAc binding present high conformational flexibility. However, the flexibility of the loops might substantially decrease in the presence of inhibitors.[17] Surprisingly, MD simulations performed by Caflisch and coworkers revealed that the conserved asparagine residue located in the binding site (e.g., Asn1168 in CREBBP, Figure 5) is able to reversibly rotate its side chain orienting it towards the solvent, especially on those bromodomains belonging to family VII in the phylogenetic tree (Figure 4).[43] Moreover, the conserved tyrosine residue (Tyr1125 in Figure 5) is also flexible and can point towards the outside of the bromodomain, especially in BRDT(1), CREBBP, BAZ2B, SMARCA4, and BRD1, which is in agreement with experimental data (PDB codes 2RFJ and 2DKW).[43] Importantly, the mentioned tyrosine movement creates a rearrangement of the protein resulting in a smaller and occluded binding site.

The large amount of structural information available for bromodomains has enabled a careful study of their aminoacid sequence and the main differences among them have been identified, enabling the rational design of bromodomain ligands. An overview of the sequence diversity of several bromodomains is given in Table 1.

The selectivity towards different bromodomains is thought to arise from length and sequence diversity within the ZA and BC loop regions that interact with residues located in close proximity to the KAc binding site.[36, 39, 44] Whereas the differences in these loop regions within the same bromodomain family are relatively small, they are significant when comparing bromodomains belonging to different families such as CREBBP and ATAD2 (Table 1). The so called hydrophobic WPF shelf in the BET family (Table 1, in red) is in fact LPF shelf in the case of CREBBP or RVF in ATAD2. The gatekeeper residue located at the bottom of the shelf also varies between isoleucine (BD1 domains of the BET family and ATAD2), valine (BD2 domains of the BET family and CREBBP, Figure 5, B) or tyrosine (GCN5, KAT2A) residues (Table 1, in blue). These differences cause changes in the size and electronics of the binding sites within different bromodomain families that could be addressed in order to design selective bromodomain ligands.

**Table 1.** Multiple sequence alignment of human BET bromodomains, BRD2, BRD3, BRD4, and BRDT, with GCN5, CREBBP and ATAD2. Highlighted in blue is the I/V to Y sequence diference in the gatekeeper residue at the bottom of the LPF shelf (Figure 5, B, C). The WPF motif (tryptophan, proline and phenylalanine shelf) is highlighted in red, and the conserved tyrosine and asparagine residues that are key for the recognition of acetylated lysines are highlighted in yellow and green respectively.[45]

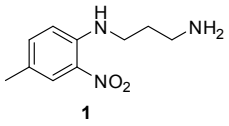
| | αZ | | | ZA loop | | αA |
|---|---|---|---|---|---|---|
| BRD2(1) | KPGRVTNQLQ | YLHKVVMKAL | WK- - -HQFAW | PFRQPVDAVK | LGLPDYHKII | KQPMDMGTIK |
| BRD3(1) | KPGRKTNQLQ | YMQNVVVKTL | WK- - -HQFAW | PFYQPVDAIK | LNLPDYHKII | KNPMDMGTIK |
| BRD4(1) | KPKRQTNQLQ | YLLRVVLKTL | WK- - -HQFAW | PFQQPVDAVK | LNLPDYYKII | KTPMDMGTIK |
| BRDT(1) | KNGRLTNQLQ | YLQKVVLKDL | WK- - - HSFSW | PFQRPVDAVK | LKLPDYYTII | KNPMDLNTIK |
| BRD2(2) | SKKGKLSEQL | KHCNGILKEL | LSKKHAAYAW | PFYKPVDASA | LGLHDYHDII | KHPMDLSTVK |
| BRD3(2) | GKKGKLSEHL | RYCDSILREM | LSKKHAAYAW | PFYKPVDAEA | LELHDYHDII | KHPMDLSTVK |
| BRD4(2) | EKSSKVSEQL | KCCSGILKEM | FAKKHAAYAW | PFYKPVDVEA | LGLHDYCDII | KHPMDMSTIK |
| BRDT(2) | VKTVKVTEQL | RHCSEILKEM | LAKKHFSYAW | PFYNPVDVNA | LGLHNYYDVV | KNPMDLGTIK |
| CREBBP | - - IFKPEELR | QALMPTLEAL | YR- - QDPESL | PFRQPVDPQL | LGIPDYFDIV | KNPMDLSTIK |
| GCN5/KAT2A | - - - - DPDQLY | TTLKNLLAQI | KS- - - HPSAW | PFMEPVKKS- | -EAPDYYEVI | RFPIDLKTMT |
| ATAD2 | QEEDTFRELR | LFLRNVTHRL | AI- - -DKRFR | VFTKPVDPD- | -EVPDYVTVI | KQPMDLSSVI |

| | -αA | αB | BC loop | | αC | |
|---|---|---|---|---|---|---|
| BRD2(1) | RRLENNYYWA | ASECMQDFNT | MFTNCYIYNK | P- - - -TDDIV | LMAQTLEKIF | LQKVASMPQE |
| BRD3(1) | KRLENNYYWS | ASECMQDFNT | MFTNCYIYNK | P- - - -TDDIV | LMAQALEKIF | LQKVAQMPQE |
| BRD4(1) | KRLENNYYWN | AQECIQDFNT | MFTNCYIYNK | P- - - -GDDIV | LMAEALEKLF | LQKINELPTE |
| BRDT(1) | KRLENKYYAK | ASECIEDFNT | MFSNCYLYNK | P- - - -GDDIV | LMAQALEKLF | MQKLSQMPQE |
| BRD2(2) | RKMENRDYRD | AQEFAADVRL | MFSNCYKYMP | P- - - -DHDVV | AMARKLQDVF | EFRYAKMPDE |
| BRD3(2) | RKMDGREYPD | AQGFAADVRL | MFSNCYKYMP | P- - - -DHEVV | AMARKLQDVF | EMRFAKMPDE |
| BRD4(2) | SKLEAREYRD | AQEFGADVRL | MFSNCYKYNP | P- - - -DHEVV | AMARKLQDVF | EMRFAKMPDE |
| BRDT(2) | EKMDNQEYKD | AYKFAADVRL | MFMNCYKYNP | P- - - -DHEVV | TMARMLQDVF | ETHFSKIPIE |
| CREBBP | RKLDTGQYQE | PWQYVDDVWL | MFNNAWLYNR | K- - - -TSRVY | KFCSKLAEVF | EQEIDPVMQS |
| GCN5/KAT2A | ERLRSRYYVT | RKLFVADLQR | VIANCREYNP | P- - - -DSEYC | RCASALEKFF | YFKLKEGGLI |
| ATAD2 | SKIDLHKYLT | VKDYLRDIDL | ICSNALEYNP | DRDPGDRLIR | HRACALRDTA | YAIIKEELDE |

In the next section, an overview of known bromodomain ligands linked to their corresponding therapeutic potential is given. Special attention is given to the BET family of bromodomains, since it has been widely explored, as well as to CREBBP and EP300, proteins of interest in this work.

### 3.1.3 Bromodomain ligands

Zhou and coworkers were the first to obtain structural information (in solution) on bromodomains in 2002, showing that PCAF bromodomain (family I, Figure 4) was able to specifically bind to HIV-1 Tat peptide upon lsyine acetylation, and that such interaction was important for HIV transcription.[30] Bromodomains were then considered as a potential therapeutic strategy against HIV.[30, 46] Zhou et al. carried out a PCAF bromodomain inhibitor discovery campaign using NMR screening techniques that resulted in compounds with affinities in the micromolar range. An example is given in Figure 6, where compound **1** is able to bind in the HIV-1 Tat peptide binding pocket with an $IC_{50}$ of 1.6 µM.




1

**Figure 6.** Structure of the PCAF inhibitor developed by Zhou et al.[46]

Since this breakthrough on the first non-peptidic small molecule acting as bromodomain inhibitor, the bromodomain field has gained increased attention and several bromodomain inhibitors have been reported in peer reviewed journals.[41-42, 47]

From the reported bromodomain inhibitors, three key interaction areas can be distinguished in the bromodomain binding site: the KAc binding pocket, the LPF/WPF shelf and the ZA channel (Figure 7). The KAc mimic has been usually achieved by methyl-substituted five-membered ring heterocycles that form hydrogen bonds with the conserved asparagine and tyrosine residues. The methyl group of the KAc, located in a small hydrophobic pocket formed by a valine and phenylalanine residues, has proven to be important for binding affinity. As all bromodomains present a highly conserved KAc binding pocket, these interactions are unlikely to be responsible for bromodomain selectivity, but more likely orient the ligand and play a role in bromodomain recognition.

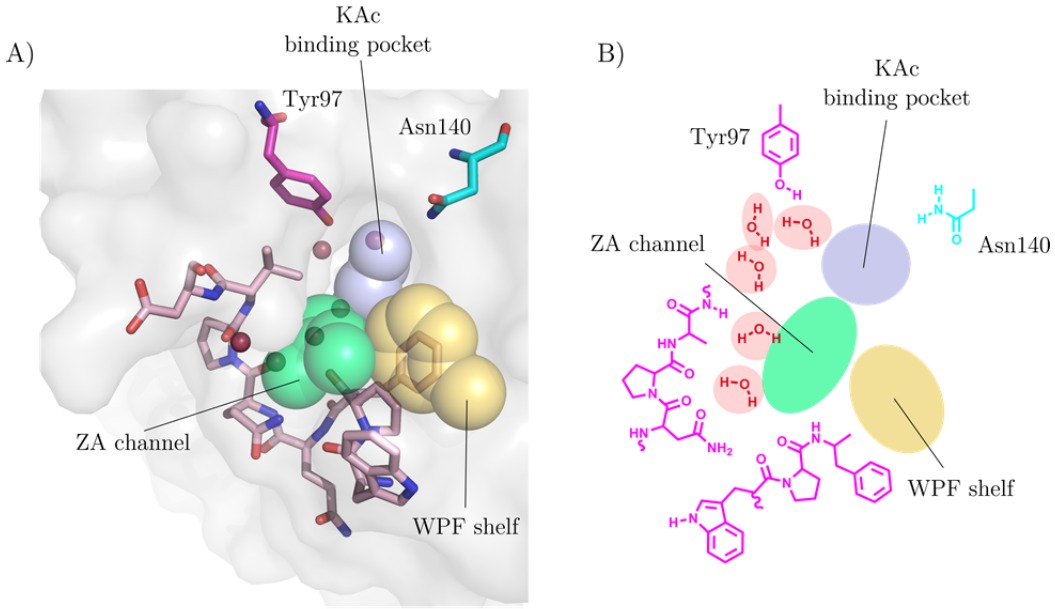



**Figure 7.** 3D (A) and 2D (B) representation of the bromodomain binding site divided into three regions: KAc binding site (in blue), WPF shelf (in yellow) and ZA channel (in green). The conserved asparagine Asn140 and tyrosine Tyr97 residues responsible for ligand binding are shown in light blue and pink respectively, together with the aminoacid residues belonging to the ZA channel and WPF shelf. The conserved crystallographic water molecules are also indicated as red spheres (A) or ellipses (B) in red. Figure recreated from PDB code 3MXF.

The LPF or WPF shelf (composed of different aminoacids according to the bromodomain, Table 1) is mainly hydrophobic and interacts with lipophilic and aromatic substituents of the ligand contributing to bromodomain selectivity and potency (Figure 7, in yellow).

The third region that can be distinguished within the binding site is the ZA channel (Figure 7, in green), which has been explored to gain potency and selectivity of small molecule ligands. There is usually one conserved water molecule within the ZA channel that forms interactions with the backbone residues of the protein and the rest of the conserved water molecules (Figure 7B, in red ellipses). All of the conserved water molecules define the shape and size of the binding site and need to be taken into account during the design of bromodomain inhibitors.

### I. BET family ligands

The BET family of bromodomains is composed of four members in mammals: BRD2, BRD3, BRD4 and BRDT (Figure 4, family II), each containing two conserved N-terminal bromodomains. Among the studied bromodomain targets, the BET bromodomain family has received preminent attention as they are known to control the transcription of virus replication (HIV[48] and EBV[49] replication in the case of BRD4)[50-51] and inflammatory responses.[52] Indeed, NF-κB, a transcription factor that regulates the immune response of the cell, can be acetylated in its Lys30 residue, causing BRD4 binding and resulting in the transcription of NF-κB genes.[52-53]

Special attention has been drawn towards BRD4 bromodomain due to its key role in several cellular processes including mitosis.[54-55] BRD4 is a key player in the regulation of transcriptional elongation by being able to recruit P-TEFb (positive transcription elongation factor).[54, 56-57] As a consequence, BRD4 has been described as an attractive target related to cancers such as acute myeloid leukaemia,[58, 17] Burkitt's lymphoma,[49] NUT midline carcinoma,[59] breast[60] and colon cancer,[61] as well as to inflamatory diseases.[52] In fact, several bromodomain ligands from the BET family of bromodomains (family II, Figure 4) are currently in clinical trials (see subsection F).[47, 62]

Most of the potent and selective bromodomain inhibitors discovered up to date target the BET bromodomain family, namely BRD2, BRD3, BRD4 and BRDT. However, the chemical diversity of BET inhibitors is quite limited and they can only be classified into a few different chemical scaffolds: methyl-triazolodiazepines and related chemotypes (**2-6**),[17, 20, 45, 63-67] 3,5-dimethylisoxazoles (**7-11**)[21, 68-73] and other chemotypes, including 3-methyldihydroquinazolinones (**12**),[74-75] 1-acyltetrahydroquinolines (**13**),[76-77] pyrazolopyridazines (**14**),[78] pyrrolopyridinones (**15**),[79-80] benzimidazoles (**16**),[81] and 2-thiazolidinones (**17**),[82] (Figure 8).

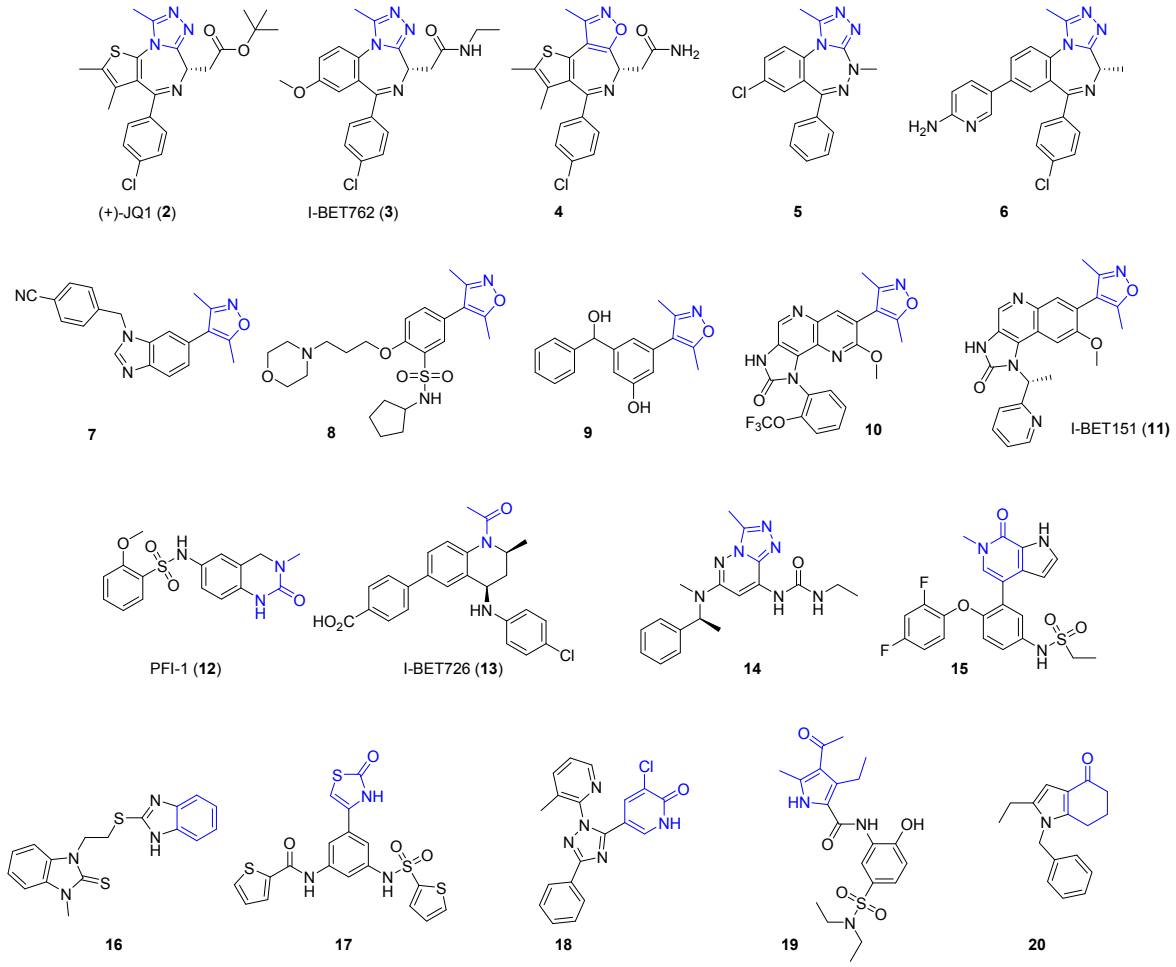



**Figure 8.** Reported chemotypes of BET bromodomain ligands. The structural motifs acting as KAc mimic are shown in blue.

In the next sections, two of the most studied scaffolds, triazolodiazepines (**2**, **3**, **5**, **6**) and dimethylisoxazoles (**7**-**11**) will be presented in detail, and other chemotypes able to mimic KAc will also be introduced (**12**-**20**).

### A) 1, 2, 4-Triazolodiazepines

The first BET potent and selective ligands, methyl triazolodiazepines **2** and **3**, were discovered in 2010 by Filippakopolus et al.[17] and Nicodeme et al. respectively.[20] These molecules present a previleged diazepine structural core that is present in central nervous system drugs such as alprazolan and triazolam. Whereas Filippakopoulos and coworkers based their drug discovery campaign in high throughput techniques and docking studies from knowledge stemming from a patent from Mitsubishi Pharmaceuticals,[83] Nicodeme et al. performed a phenotypic screening to discover up-regulators of apolipoprotein A1 (ApoA1), which is involved in protection from atherosclerosis and anti-inflammation.[84]

(+)-JQ1 (**2**) is a potent BRD4(1) inhibitor ($K_d$ = 49 nM determined by ITC, Isothermal Titration Calorimetry, and $IC_{50}$ = 77 nM determined by AlphaScreen) and is selective against

the BET family of bromodomains (family II, Figure 4) according to differential scanning fluorimetry (DSF).[17] (+)-JQ1 (**2**) is able to displace BRD4 from chromatin in cell based assays using fluorescence recovery after photobleaching (FRAP). Furthermore, it inhibits the proliferation of various cancer cells and its *in vivo* efficacy has been demonstrated on a NUT midline carcinoma (NMC) mice xenograft. Further studies have also shown efficacy in metabolic disease[85] and acute myeloid leukemia mice models[86] and non-small cell lung cancer cells.[87]

I-BET762 (**3**) shows similar potency as (+)-JQ1 (**2**) with a $K_d$ of 55 nM (measured by ITC) and an $IC_{50}$ of 36 nM according to a FRET assay against BRD4(1).[45] Compound **3** is able to suppress the expression of pro-inflammatory genes *in vivo* and inhibits the proliferation of several cancer cells.[45] Its efficacy has also been demonstrated in several tumor and immune-inflammatory models, and is currently undergoing clinical phase I/II studies (see sub-section F).[61, 71, 88]

From the structure activity relationship (SAR) studies performed by both research groups the importance of the absolute configuration at the molecule's stereocenter can be highlighted, as the opposite enantiomers were inactive due to steric clashes with residues of the ZA loop region. The crystal structures of **2** and **3** have been solved with a variety of BET bromodomains showing excellent shape complementarity with the KAc binding site thanks to the curvature of the triazolodiazepine core. Both ligands present almost identical binding poses and occupy the totality of the binding site. The key interactions of (+)-JQ1 (**2**) and I-BET762 (**3**) in BRD4(1) are shown in Figure 9.

The triazole ring of (+)-JQ1 (**1**) and I-BET762 (**3**) is deeply inserted into the binding site acting as a KAc mimic, with the methyl group pointing into the small hydrophobic pocket that recognizes the methyl group of KAc (Figure 9). SAR studies performed by Filippakopoulos et al. proved the methyl group to be essential for binding affinity, as its replacement by other groups (H, Et, Ph, $NH_2$ or carbonyl) was detrimental for binding.[17] One of the nitrogen atoms of the triazole ring accepts a water mediated hydrogen bond from the evolutionary conserved tyrosine (Tyr97) and the second nitrogen atom accepts a hydrogen bond from the conserved asparagine Asn140 (Figure 9 B, D, hydrogen bonds are shown as dashed lines).[20, 45]

The 4-chlorophenyl (**2-4**, **6**) and phenyl ring (**5**) moieties, common for several analogues,[45] protrude out of the binding pocket and interact in the hydrophobic WPF shelf, which is composed of residues Trp81, Pro82 and Phe83 located in the ZA loop (Figure 9). The fused aromatic ring of the triazolodiazepine core (the dimethylthiophene in compound **2** and methoxybenzo moiety in compound **3**) are located in the lipophilic ZA channel and mainly establish van der Waals interactions (Figure 9).

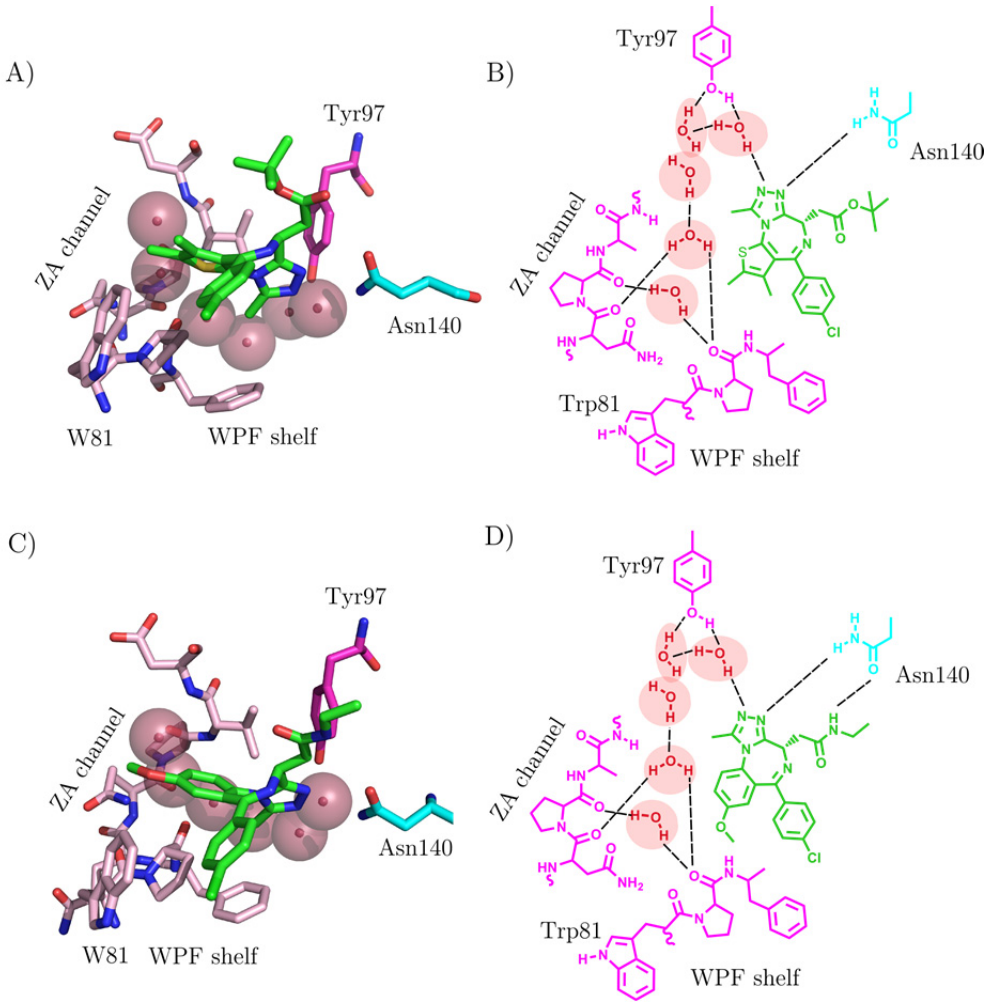



**Figure 9.** A, C) Molecular interactions of (+)-JQ1 (**2**) (A, PDB code 3MXF) and I-BET762 (**3**) (C, PDB code 3P5O) with BRD4(1). B, D) The acetyl-lysine binding pocket of BRD4(1) is shown as a 2D representation in complex with ligands **2** and **3**. The inhibitors are shown in green. The conserved residues tyrosine Tyr97 and asparagine Asn140 responsible for binding and the residues belonging to the WPF shelf and ZA channel are indicated. The hydrogen bonds between the protein and the inhibitors are indicated as dashed lines.

### B) 3,5-Dimethylisoxazoles

3,5-Dimethylisoxazoles have been independently identified as BRD4(1) inhibitors by Dawson et al.,[21] Hewings et al.[68] and Bamborough et al.[69]

Hewings and coworkers discovered that the solvent N-methyl pyrrolidinone (NMP) binds weakly within the KAc binding pocket, and a subsequent screening of methyl-bearing heterocycles resulted in dimethylisoxazole **21** as a hit.[68] Starting from the dimethylisoxazole hit **21**, Hewings et al. developed lower molecular weight fragments resulting in compounds **22** and the most potent compound **9** ($IC_{50}$ = 382 nM according to AlphaScreen), as illustrated in Figure 10.[68] The dimethylisoxazole ring binds within the KAc-binding pocket as a KAc isostere: the oxygen atom of the five membered ring (mimic of the carbonyl oxygen of KAc) accepts a

hydrogen bond from the conserved asparagine Asn140 (Figure 11B, hydrogen bonds shown as dashed lines), the nitrogen acts as a hydrogen bond acceptor from tyrosine Tyr97 (Figure 11B, hydrogen bonds shown as dashed lines) and one of the methyl groups occupies a small hydrophobic pocket analogous to the terminal methyl in KAc.

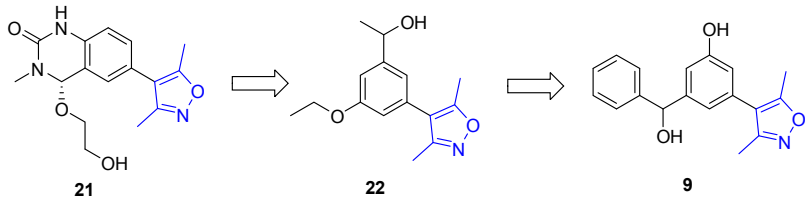



**Figure 10.** Starting hit (**21**) and optimized dimethylisoxazole derivatives (**22** and **9**) as BET inhibitors reported by Hewings et al.[68, 89] The structural motifs acting as KAc mimic are shown in blue.

The hydroxyl group in the phenol moiety is involved in hydrogen bonds with the conserved crystallographic water molecules within the ZA channel, whereas the hydroxy group attached to the stereogenic center in **9** interacts with the solvent, making its configuration not relevant for binding (Figure 11). Finally, the phenyl substituent lies within the WPF shelf, in a similar region as the chlorophenyl in tetrazolodiazepines **2** and **3** (Figure 9).[17, 20] This compound series presents higher ligand efficiency values than the triazolodiazepine inhibitors (Figure 8, top) due to their low molecular weights, making them promising leads for the development of more potent BET inhibitors. Moreover, sub-microlar $GI_{50}$ values were obtained for compound **9** against an acute myeloid leukemia cell line.[68, 89]

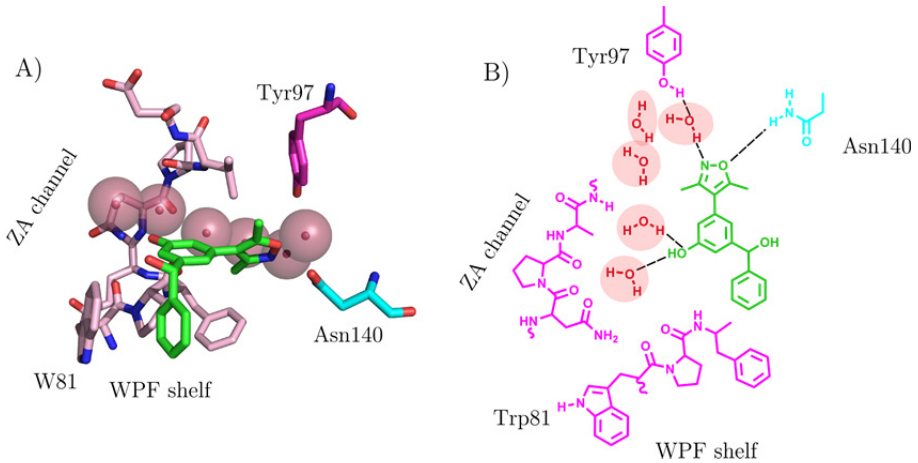



**Figure 11.** A) Molecular interactions of **9** (PDB code 4J0S) with BRD4(1). B) The acetyl-lysine binding pocket of BRD4(1) is shown as a 2D representation in complex with ligand **9**. The inhibitor is shown in green. The conserved residues tyrosine Tyr97 and asparagine Asn140 responsible for binding and the residues belonging to the WPF shelf and ZA channel are indicated. The hydrogen bonds between the protein and the inhibitor are indicated as dashed lines.

Independently, Bamborough and coworkers screened potential fragments as KAc mimics towards BET bromodomains and found low molecular weight 3,5-dimethylisoxazoles as good candidates.[69] Compounds bearing a sulfonamide group on the phenyl ring in *meta* relative position to the isoxazole presented higher binding affinities, as the sulfonamide introduces a bent in the shape of the molecule directing the cycloalkyl substituents into the WPF shelf. Introduction of polar and solubilizing groups in the molecule resulted in isoxazole **8**, a molecule that shows high selectivity for the BET bromodomains and low micromolar $IC_{50}$ values on a cell based cytokine IL-6 assay.[69]

Another example of 3,5-dimethylisoxazoles was discovered by a phenotypic screening campaign in search for ApoA1 up-regulators by Dowson and coworkers.[21] A medicinal chemistry campaign yielded I-BET151 (**11**), which presents high affinity towards BRD3 *in vitro* and nanomolar $IC_{50}$ values towards MLL fusion leukemia cell lines. Its therapeutic potential has also been established in a MLL leukemia *in vivo* model.[21]

Aside from the above-mentioned ligands, dimethylisoxazoles have also been greatly explored by other groups leading to the development substituents with different chemotypes as illustrated in Figure 8 (**7**,**10**).[70, 73]

### C) Other chemotypes

Even though triazolodiazepines and dimethylisoxazoles have been two of the most studied scaffolds within the bromodomain inhibitor pool, a few additional chemotypes have been identified as bromodomain ligands (Figure 8, **12**-**13**). Their binding interactions within the KAc binding pocket are illustrated in Figure 12.

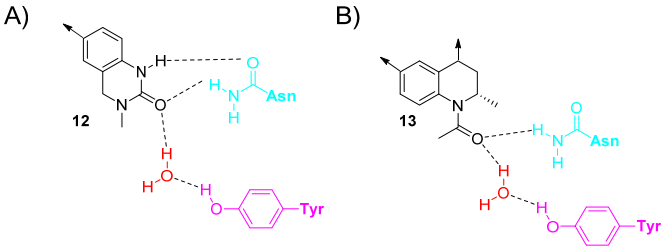



**Figure 12.** Schematic representation of the binding modes of the KAc mimics of compounds **12** and **13** in the binding site.[90] The arrows indicate the connection vectors with the rest of the molecule. The hydrogen bonds between the protein and the inhibitors are indicated as dashed lines.

The dihydroquinazoline-3,5-dimethylisoxazole **21** (Figure 10)[68] presents two distinct binding modes. In both, the 3,5-dimethylisoxazole or the dihydroquinazoline can act as KAc mimic. Taking the dual binding mode of compound **21** into account, Knapp and coworkers explored the utility of the dihydroquinazoline core as BET bromodomain inhibitor.[74] The study led to the discovery of benzenesulfonamide-quinazolin-2-ones, being PFI-1 (**12**) one of the most

potent derivatives found in this campaign. The cyclic urea of the dihydroquinazoline ring is buried in the binding pocket mimicking the KAc ligand (Figure 12, A). In an analogous approach to the work of Bamborough and coworkers in the design of isoxazole **8**,[69] sulfonamide linkers were incorporated to direct its aromatic substituents towards the WPF shelf.

PFI-1 (**12**) shows high specificity towards the BET bromodomain family, even though a certain binding towards CREBBP and EP300 is also observed by differential scanning fluorimetry (DSF).[74] Its cellular activity was confirmed in an IL-6 assay (a cytokine that is produced by NF-κB) and antiproliferative activities with $GI_{50}$ values in the low micromolar level were observed in leukemia cell lines. Moreover, PFI-1 (**12**) shows a promising pharmacokinetic profile *in vivo* for oral administration.[74-75, 91]

A number of 1-acyltetrahydroquinoline derivatives have been patented by GlaxoSmithKline, including I-BET762 (**13**), as KAc mimics (see Figure 12, B for an schematic representation of its binding mode).[76] Compound **13** is able to inhibit the growth of several cancer cell lines and its efficacy *in vivo* for treating neuroblastoma and endotoxemia has been confirmed.[76, 92]

Further examples of BET bromodomain ligands have been published as patents, including pyrazolopyridazines (**14**) by Costellation Pharmaceuticals[78] and pyrrolopyridinones and pyrrolopyridazinones (**15**) by GlaxoSmithKline, Constellation Pharmaceuticals, Pfizer, Dana-Farver Cancer Institute and Abbott Laboratories.[79-80]

### D) Virtual screening as a tool for the discovery of novel chemotypes

Virtual screening has been a very efficient tool in the discovery of novel KAc mimetics as bromodomain ligands (Figure 13).[81-82, 93-95]

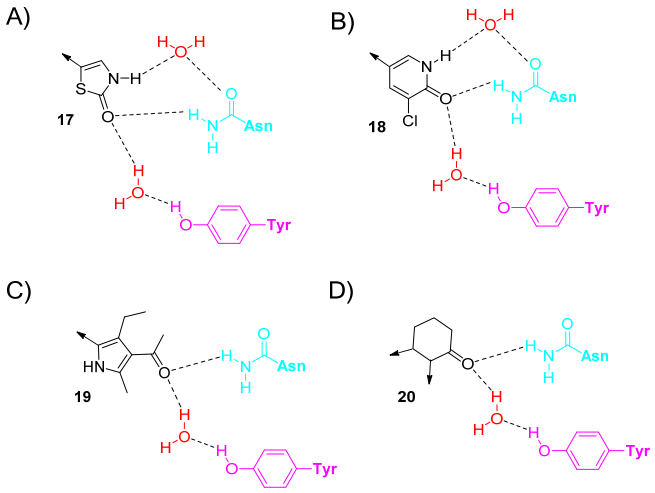



**Figure 13.** Schematic representation of the binding modes of the KAc mimics of compounds **17-20** in the binding site.[90] The arrows indicate the connection vectors with the rest of the molecule. The hydrogen bonds between the protein and the inhibitors are indicated as dashed lines.

The first *in silico* guided design of bromodomain inhibitors was conducted by Ito and coworkers in 2011. An initial high throughput docking campaign on BRD2(1), followed by the screening of 192 compounds experimentally afforded benzimidazole **16** (Figure 8), which was able to inhibit BRD2(1) with a $K_d$ value of 28 µM.[81] Unexpectedly, the benzene ring occupies the KAc binding site without forming any interactions with the conserved asparagine Asn140 and tyrosine residue Tyr97, which also explains the low affinity of the ligand (see PDB code 3AQA for further details).

Another virtual screening campaign performed in 2013 by Shen and coworkers resulted in the identification of new chemotypes (e.g. thiazolidinones) as potential BET inhibitors.[82] After docking a subset of the ZINC library with Glide on BRD4(1), 41 fragments were selected, and X-ray crystallography was used to further identify hits. The 2-thiazolidinone moiety was found to act as a KAc mimic (Figure 13, A). Based on the crystal structure (PDB code 4HXN), the initial 2-thiazolidinone fragment was further optimized, including, once more, sulfonamide substituents that resulted in an increase of potency, being thiazolidinone **17** (Figure 8) the most active analogue with an $IC_{50}$ value of 0.23 µM in BRD4(1).[82]

In the same year, Hoelder and coworkers were able to identify new scaffolds as KAc mimetics by a virtual screening approach against BRD4(1).[93] Commercially available compounds that matched the hydrogen-bond pattern and steric constraints of the KAc binding site were selected and docked with the software Glide, resulting in unprecedented KAc mimetics with $IC_{50}$ values in the µM range: chloropyridones (being compound **18** the most active from the series, Figure 13, B), triazolopyrimidines, quinolones, pyrrolopyridones and pyrrolopyrimidones.[93]

Few months later, a new *in silico* screening campaign against BRD4(1) was carried out by Günther and coworkers.[94] A library composed of 7 million molecules was screened and 22 compounds were selected and purchased, leading to seven compounds with significant binding affinity (six of which were unkown scaffolds). The 4-acyl pyrrole **19** (Figure 7, PDB code 3UVW), with a $K_d$ of 237 nM and a $GI_{50}$ value of 17 µM in a leukemia cell line, constitutes an interesting new scaffold for the design of potent BET inhibitors.[94] The pyrrole ring binds within the KAc binding pocket showing high surface complementarity, while the acyl substituent interacts with the conserved asparagine and tyrosine residues (Figure 13, C).

Our collaborators, the group of Prof. Caflisch in the Biochemistry Institute (UZH) have greatly contributed to the bromodomain research in the last year.[43, 95] In fact, they implemented an *in house* developed fragment based *in silico* ALTA procedure[96] for the discovery of new BRD4(1) ligands, resulting in the low micromolar active cyclohexanone **20** (Figure 8) among other active chemotypes.[95] The binding mode of the KAc mimetic of compound **20** is shown in Figure 13, D (PDB code 4PCE).

### E) Kinase inhibitors with bromodomain activity

Interestingly, two similar studies published almost simultaneously identified kinase inhibitors (some of them clinically approved) as BRD4(1) nanomolar ligands.[97-98]

Ember et al. assessed the binding of more than 500 known kinase inhibitors by robotic co-crystallization on BRD4(1) and BRDT(1) obtaining X-ray structures of 14 kinase inhibitors in complex with bromodomains.[98] Independently, the Structural Genomics Consortium (SGC), in collaboration with DiscoveRx, screened 628 kinase inhibitors against BRD4(1) using an AlphaScreen assay.[97] Surprisingly, both studies led to kinase inhibitors BI-2536 (**23**) and TG-101348 (**24**) as two of the most potent compounds against BRD4(1), with $K_d$ values of 37 and 164 nM respectively and a high selectivity towards BET bromodomains (Figure 14).[97-98] Both compounds were able to displace BRD4 from chromatin in cells, as well as to suppress c-Myc oncogene expression (a marker of BET inhibition).[97]

These results suggest that kinases and bromodomains share a high degree of pharmacophore similarity. However, the chemical moieties of the kinase inhibitors that interact as KAc mimic do not always represent the groups interacting with the ATP-binding site of the kinase. The fact that BET bromodomains might be off-targets of kinase inhibitors constitutes a very important piece of information that will very likely impact bromodomain and kinase research (see chapter 2 for more information on kinase inhibitors).

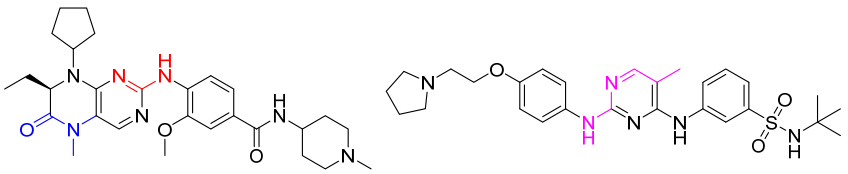



**Figure 14.** Two kinase inhibitors, **23** and **24** show potent BRD4(1) activity. The atoms highlighted in blue in **23** correspond to the KAc mimic for bromodomain binding and those in red interact with the hinge region of the kinase (see chapter 2 for further details). In TG-101348 (**24**), the same moiety is responsible for both, bromodomain and kinase binding (indicated in pink).

### F) Clinical studies with BET bromodomain ligands[41, 47]

Based on all the BET bromodomain ligands available in the literature, we can conclude that numerous efforts have been invested in the last 3-4 years in the development of BET family bromodomain ligands, both in industry and academia. The high number of X-ray crystal structures has allowed a thorough study of the potency and selectivity of protein-ligand interactions. However, the usefulness of bromodomain ligands for therapeutical applications remains an open question.[47] The five BET bromodomains that are currently undergoing clinical trials might shed some light into the validation of bromodomains as therapeutical targets.[62]

The clinical candidates for which their chemical structure has been disclosed are shown in Figure 15. I-BET762 (**3**) is being pursued against NUT midline carcinoma (NMC), solid tumors and haematological malignancies, RVX-208 (**25**) towards the treatment of atherosclerosis and type II diabetes and OTX015 (**26**) against acute leukemia and hematological malignancies. Other clinical candidates include CPI-0610 (for the treatment of progressive lymphoma) and TEN-010, which is a derivative of (+)JQ1 (**2**) and is being evaluated against NMC.

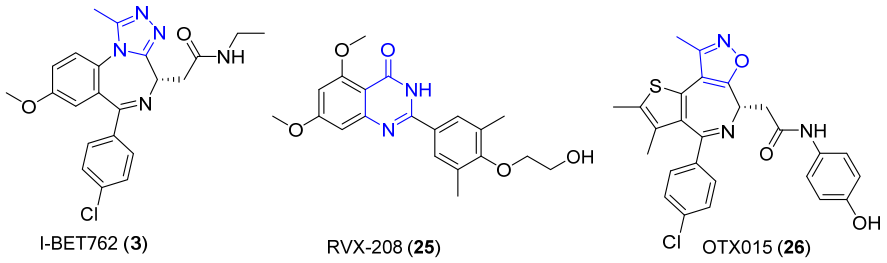



**Figure 15.** Bromodomain ligands currently in clinical trials.[62] The structural motifs acting as KAc mimic are shown in blue.

### II. CREBBP and EP300 ligands

The bromodomain containing protein CREBBP is gaining more attention as a target for the development of small organic inhibitors. EP300 and CREBBP belong to the same bromodomain family (family III, Figure 4) and share 96 % sequence similarity.[99] They are part of a multidomain protein that contains HAT (histone acetyltransferase) catalytic activity, meaning that they are able to acetylate conserved lysine amino acids on histone or non-histone proteins, as well as having an acetylated lysine binding domain (bromodomain).[100-101] Due to the acetyltransferase and bromodomain activities, CREBBP and EP300 have at least 400 interacting protein partners (such as transcription factors and co-activators and acetylated histones),[102] which highlights their importance within the protein-protein interactome.[103]

CREBBP and EP300 play important roles in DNA replication and repair, cell growth and cell cycle regulation and genomic stability.[104-105] As an example, EP300 and CREBBP are able to acetylate p53 in its K382 residue upon extracellular stress or DNA damage through their HAT domain and they are also known to specifically bind to acetylated p53 via their bromodomain.[106] As a consequence, changes in the p53-dependent activation of target genes take place, resulting in cell cycle arrest, senescence or apoptosis. [107-109] p53 is a tumor suppressor gene, and as such, its mutations are known to occur in about 50 % of human cancers.[107, 110-111] In addition, p53 is implicated in a wide variety of malignancies such as Alzheimer's disease, Parkinson's disease, spinal cord diseases, ischemic brain injury, infectious and auto-immune related disorders (including multiple sclerosis) and myocardial ischemia.[99, 112-117]

Interestingly, CREBBP and EP300 have been described to have opposite effects in oncology, being either oncogenes or tumor supressors.[102, 118] On the one hand, chromosome

translocations resulting in gene fusions containing CREBBP or EP300 have been linked to leukemias and lymphomas.[102, 119] On the other hand, CREBBP and EP300 are mutated in solid tumors and B cell lymphoma, suggesting they possess a tumor suppressing role.[105, 120]

CREBBP and EP300 were first targeted to inhibit their acetyltransferase activity.[102, 121-122] The inhibitors suppressed lysine acetylation and cell proliferation, and their efficacy was also demonstrated in *in vivo* cancer xenograft models.[102, 121-122] Nowadays, there is an emerging strategy to target the CREBBP and EP300 bromodomains, which has resulted in the development of a few CREBBP and EP300 ligands in the last year.

The main structural difference between the BET family and CREBBP bromodomain lies within the hydrophobic LPF shelf (Figure 16). The residue Trp81 of the WPF shelf in BRD4(1) (Figure 16, A) is substituted by Leu1109 in CREBBP (Figure 16, B). As Leu1109 is located further away from the binding site compared to Trp81, a larger space arises in CREBBP (Figure 16, indicated as pink spheres). More importantly, the presence of the arginine residue Arg1173 (also missing in the BET bromodomain family) provides a basic environment in the entrance of the binding site in CREBBP. Those two main differences are thought to contribute towards CREBBP selectivity and have been addressed in the development of CREBBP inhibitors.[68, 99, 123-125]

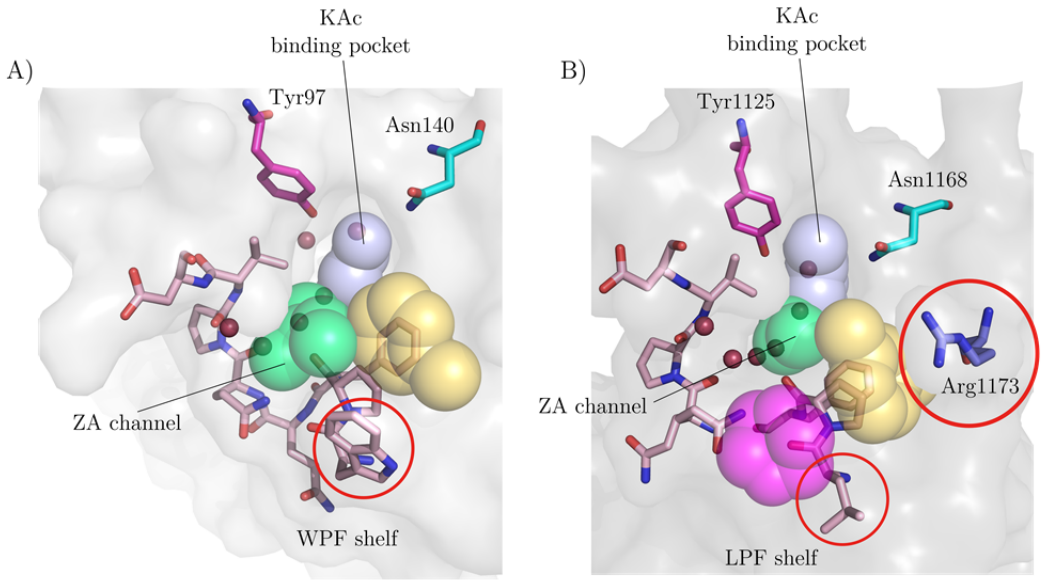



**Figure 16.** 3D representation of a BET bromodomain (A, PDB code 3MXF) and CREBBP bromodomain (B, PDB code 4NR7) binding site divided into three regions: KAc binding site (in blue), WPF or LPF shelf (in yellow) and ZA channel (in green). The main differences between both bromodomains are highlighted by a red circle.

The first CREBBP ligands were developed by Zhou and coworkers in 2006 via NMR screening of 200 compounds that contained a KAc mimic.[123] As the entrance to the CREBBP bromodomain is positively charged due to the arginine residue Arg1173, the compound library

was composed of electron-rich molecules. The study resulted in fourteen ligands including compounds **27** and **28**, which were able to block the interaction between p53 KAc382 and CREBBP (Figure 17).

Following a similar approach, the same group disclosed azobenzenes as potential bromodomain inhibitors.[124] An optimization campaign yielded ischemin **29**, which has a $K_d$ value of 19 µM and an $IC_{50}$ value of 5 µM in a cell based p21 luciferase assay (a p53 target gene).[124] The NMR structure of compound **29** shows that the ligand does not bind deep within the binding site (Figure 18) and it only forms one hydrogen bond with the conserved asparagine Asn1168 via the phenol. On the other end of the molecule, the sulfonate substituent forms an electrostatic interaction with arginine Arg1173, located at the entrance of the binding site. The same arginine residue also interacts with the aniline through hydrogen bonding (Figure 18, B, hydrogen bonds shown as dashed lines). Ischemin has moderate selectivity: it is 5-fold more selective for CREBBP over PCAF, BRD4(1) and BAZ2B based on a fluorescence assay.

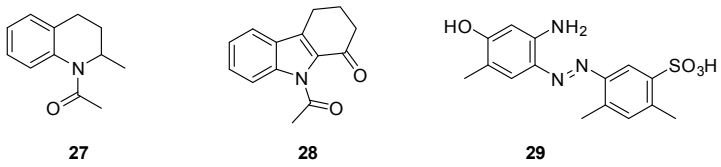



**Figure 17.** First CREBBP ligands developed by Zhou et al.[123-124]

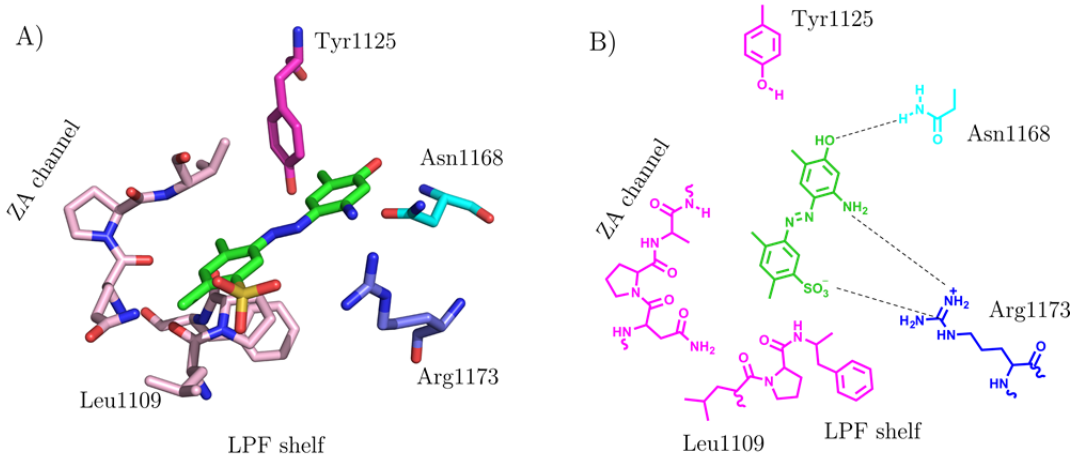



**Figure 18.** A) Molecular interactions of **29** (PDB code UL84) with CREBBP. B) The acetyl-lysine binding pocket of CREBBP is shown as 2D representation in complex with ligand **29**. The inhibitor is shown in green. The conserved residues tyrosine Tyr1125 and asparagine Asn1168 responsible for binding and the residues belonging to the LPF shelf and ZA channel are indicated, together with Arg1173, a characteristic residue of CREBBP that is located in the BC loop. The hydrogen bonds between the protein and the inhibitors are indicated as dashed lines.

A few medicinal chemistry campaigns in the course of developing BET bromodomain ligands (see sub-section I) have also yielded CREBBP ligands with moderate to low selectivity. Two examples are shown in Figure 19.

Out of the numerous 3,5-dimethylisoxazol derivatives developed by Hewings et al.[63] (see sub-section I), compound **30** has an $IC_{50}$ value of 32 µM and 28 µM against CREBBP and BRD2(1) respectively.[63] Interestingly, the carboxylate group mimics the sulfonate in the azobenzene **29** (Figure 17) by interacting with arginine Arg1173, an interaction that might be responsible for CREBBP activity (for further details see PDB code 3SVH).

Another example of non-selective CREBBP ligands is 1,2,4-triazolo[4,3-a]phthalazine **31**, which was discovered by Brennan et al. in 2013.[126] The phthalazine derivatives maintained the 3-methyl-1,2,4-triazolo core of potent BET inhibitors such as (+)-JQ1 (**2**) and I-BET762 (**3**) and varied the fused ring and substituents present in the ZA channel and WPF shelf in order to change the selectivity towards bromodomains outside the BET family. After the addition of solubilizing groups, the chloro substituted sulfonamide **31** (Figure 19) was found to be potent towards CREBBP ($IC_{50}$ = 0.20 µM), although showing, at the same time, a higher potency against BRD4(1) with an $IC_{50}$ of 0.16 µM.[126]

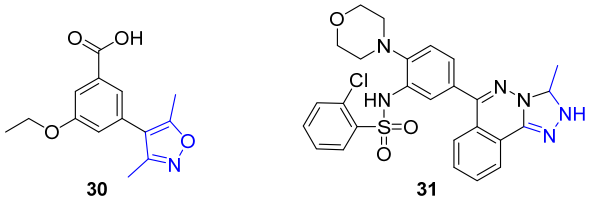



**Figure 19.** CREBBP inhibitors steaming from BET inhibitor optimization campaigns.[68, 126] The structural motifs acting as KAc mimic are shown in blue.

The first CREBBP potent and selective ligands were recently and almost simultaneously reported by the group of Conway[125] and the Structural Genomics Consortium (SGC)[99, 127-128] (**32-34**, Figure 20). Both studies originated from KAc mimics already known for BET inhibitors (Figure 8).

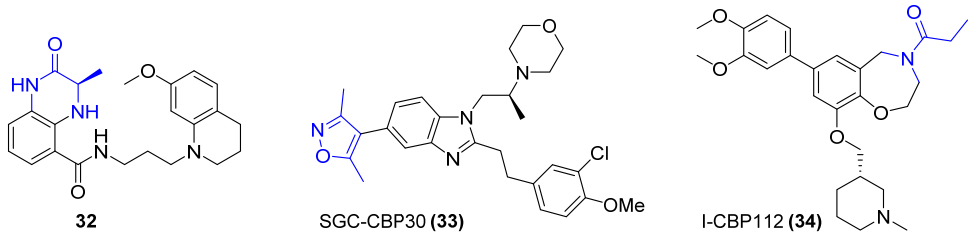



**Figure 20.** First potent and selective CREBBP inhibitors.[99, 125, 127-128] The structural motifs acting as KAc mimic are shown in blue.

Conway and co-workers initiated their study inspired by the dihydroquinazoline KAc mimic already explored as BET ligands (compound **12**, Figure 8).[125] Since the dihydroquinazoline motif was found to be susceptible to oxidation,[68, 89] more stable KAc mimics were investigated, including benzoxazinones and dihydroquinoxalinones, the latter resulting in more potent ligands. The amide of the dihydroquinoxalinone forms two hydrogen bonds with Asn1168, while the methyl group occupies the same small hydrophobic pocket as the acetyl gorup in KAc (Figure 21, A, B). The intramolecular hydrogen bond formed between the secondary amine and the carbonyl of the amide provides the right orientation for the tetrahydroquinoline substituent to reach an induced fit pocket below the Arg1173 residue establishing a cation-π interaction (Figure 21, A, B).[125] An optimization of the mentioned interaction with Arg1173 by introducing electron rich substituents in the aromatic ring, lead to the most potent ligand in the series, **32**, which shows high CREBBP potency (an $IC_{50}$ value of 323 nM by Alphascreening and a $K_d$ of 390 nM determined by ITC) and a 3.5 fold selectivity over BRD4(1) (with a $K_d$ of 1.4 µM).[125] Compound **32** was also able to displace the CREBBP bromodomain from chromatin in cells, as demostrated in a FRAP assay.[125]

Hay et al. developed slightly more potent and selective CREBBP ligands starting from the BET ligand isoxazole **7** (Figure 8).[99] Analogues were designed based on X-ray structures of an intermediate compound on both, BRD4(1) and CREBBP, with the aim to gain selectivity towards CREBBP. SAR studies yielded compound SGC-CBP30 (**31**), which shows an unprecedented selectivity of 40 fold for the CREBBP bromodomain against BRD4(1) (with a $K_d$ of 21 nM and 850 nM in CREBBP and BRD4(1) respectively). Due to the high similarity between CREBBP and EP300, isoxazole **33** is also able to strongly bind to EP300 ($K_d$ = 32 nM).[99]

The X-ray structure of compound **33** is shown in Figure 21, where the isoxazole acts as a KAc mimic as explained in the previous sub-section. The aryl ring bearing electron rich substituents establishes a cation-π interaction with the guanidino group of Arg1173, in analogy to compound **32**. The chiral methyl group is located below the aryl ring locking its position and providing the right orientation for the cation-π interaction to take place. Finally, the morpholino group points towards the solvent resulting in an improved water solubility, an important property when aiming for compounds that could reach the drug market. The cellular activity of isoxazole **33** was evaluated using a fluorescence recovery after photo-bleaching (FRAP) assay, concluding that the CREBBP bromodomain could succesfully be displaced from chromatin at low concentrations of the compound. Moreover, it could inhibit p53 activity in a dose dependent manner and displayed a $GI_{50}$ value of 8.1 µM against U2OS osteosarcoma cells.[99]

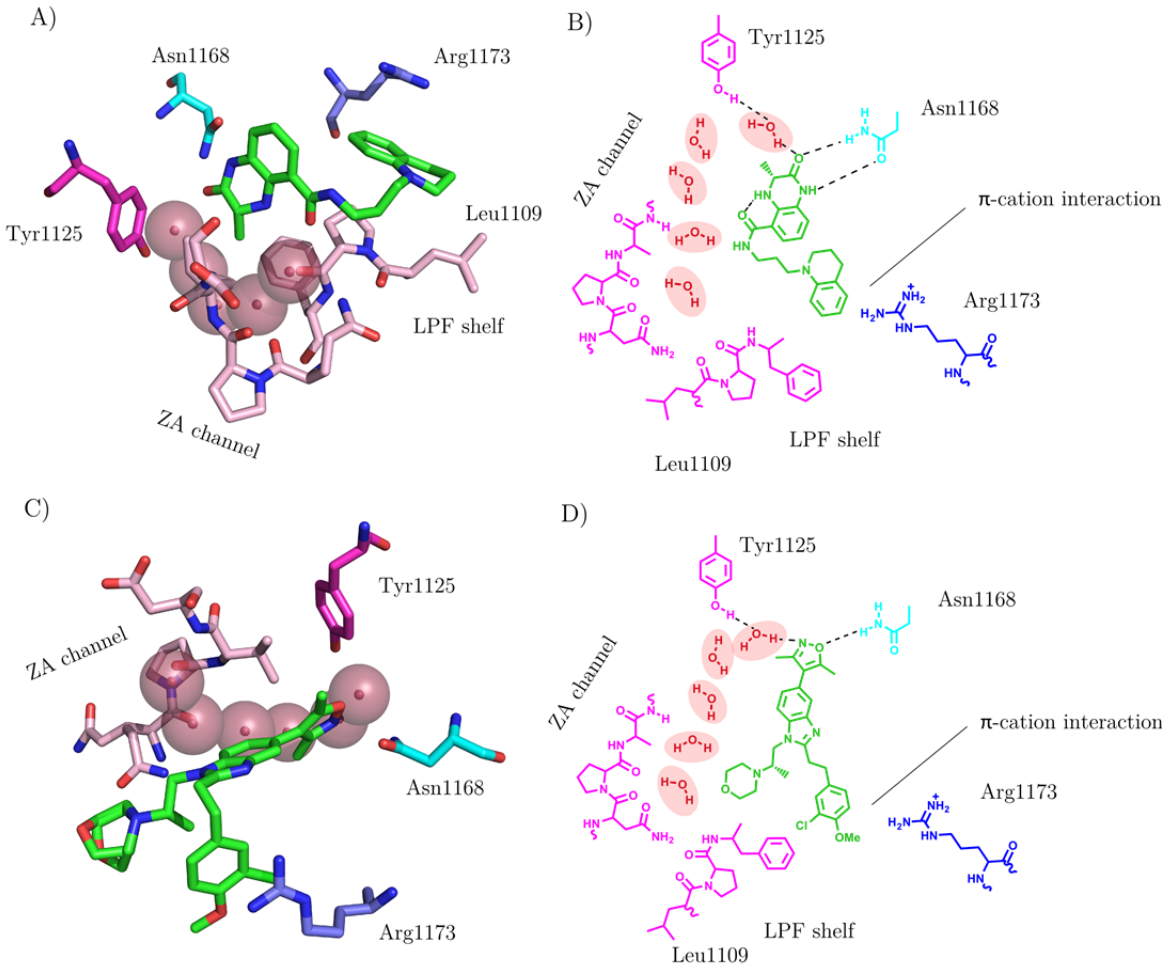



**Figure 21.** A, C) Molecular interactions of **32** (A, PDB code 4NYW) and **33** (C, PDB code 4NR7) with CREBBP. B, D) The acetyl-lysine binding pocket of CREBBP is shown as 2D representation in complex with ligands **32** and **33**. The inhibitor is shown in green. The conserved residues tyrosine Tyr1125 and asparagine Asn1168 responsible for binding together with Arg1173, a characteristic residue of CREBBP and residues belonging to the LPF shelf and ZA channel are indicated. The hydrogen bonds between the protein and the inhibitors are indicated as dashed lines.

The last known CREBBP selective inhibitor, I-CBP112 (**34**), has also been developed in the SGC.[128] It presents an $IC_{50}$ value of 170 nM by AlphaScreen and its selectivity has been confirmed against several other bromodomains by surface plasmon resonance.

Due to the scarce number of selective CREBBP ligands available, all of them, **32-34**, reported during the course of this work, we decided to initiate an *in silico* driven screening campaign, the aim being the development of novel, potent and selective CREBBP ligands. Such molecules could contribute to unravel the biological processes behind this bromodomain and find its direct therapeutical application. The results are presented in section 3.2.

### III. Other bromodomain ligands

There has been very little work done outside the BET and CREBBP/EP300 bromodomains. In fact, only three bromodomains have been addressed by potent inhibitors (Figure 22, **35-38**): BAZ2B, BRPF1 and SMARCA4, which have been characterized as non-druggable bromodomains in a study conducted by Vidler et al.[129] The biological function of the mentioned bromodomains remains elusive, but the existence of potent inhibitors might help unravel their therapeutical applications.

Importantly, bromosporine (**39**)[130] is a special tool within the bromodomain inhibitors pool that is able to bind to a diverse set of bromodomains (including BRD2, BRD3, BRD4, BRDT, PB1, CECR2, TAF1, BAZ2A, TAF1, BRD9 and CREBBP). It was developed at the SGC and could be very useful for studying the biological role of a diverse set of bromodomains as well as for validating functional assays.

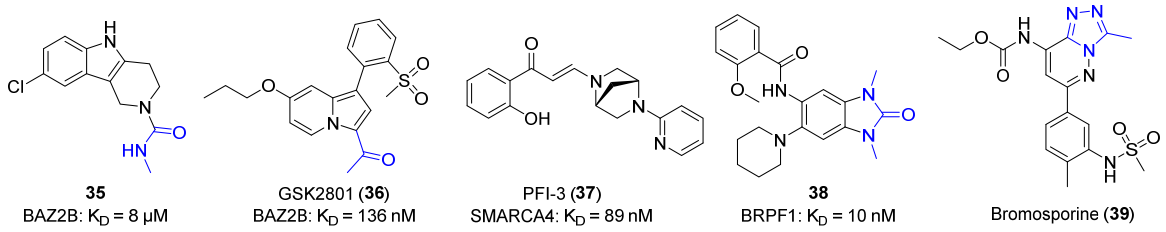



**Figure 22.** Potent inhibitors that target bromodomains outside the BET family or CREBBP/EP300 bromodomains.[131-134] The structural motifs acting as KAc mimic are shown in blue.

## 3.2 Fragment-based Design of Potent and Selective CREBBP Bromodomain Ligands: Polar Interactions with Arg1173 as a Tool Towards Selectivity

The quest for bromodomain inhibitors as potential therapeutic tools has bloomed in recent years based on the growing understanding of epigenetic processes. Earlier this year, Rooney et. al reported the first nM potent CREBBP inhibitors based on dihydroquinoxalinone scaffolds.[125] Absolute stereocontrol was required to attain the desired potency and still only moderate selectivity against BRD4(1), the most promiscuous bromodomain, was found. Recently, Hay et. al reported a medicinal chemistry optimization campaign for selective CREBBP bromodomain inhibitors starting from a non-selective 3,5-dimethylisoxazole ligand.[99]

We decided to take a different approach specifically focussing on the target[95, 135] and started our CREBBP ligand-identification campaign by fragment-based high-throughput

docking into the structure of the CREBBP bromodomain. Herein we present the result of a computer-aided, structure-based approach which has enabled the discovery of several nM ligands of the CREBBP bromodomain upon optimization of binding by modulation of the electrostatic interaction with Arg1173 (numbering from PDB structure 3SVH), a residue located at the entrance of the binding site that is considered to be key for attaining selectivity towards CREBBP.[68, 123-124] The anchor-based library tailoring procedure (ALTA)[136-138] was used for high-throughput docking (further details can be found in the experimental section). Initially, the nearly two million compounds in the ZINC leads-now library (version of Oct. 2012) were decomposed into fragments by the program DAIM.[139] Upon filtering out fragments without rings and hydrogen bond donor/acceptor groups, a total of 106,534 fragments were docked by the program SEED[140-141] which takes into account electrostatic solvation by numerical solution of the generalized Born equation.[142] The top (about 4,000) fragments were used to retrieve nearly 800,000 molecules in the ZINC library which they originated from. Docking of these molecules was carried out by AutoDock Vina[143] and the poses were ranked by the SEED energy[140-141] followed by clustering. Finally, 17 compounds (cluster representatives) were chosen for *in vitro* validation. The acyl benzene **A** showed activity and was selected for further investigations based on its amenability to potential chemical editing. Three additional acyl-benzene derivatives (compounds **B**, **C** and **D**) were also purchased and tested (Figure 23).

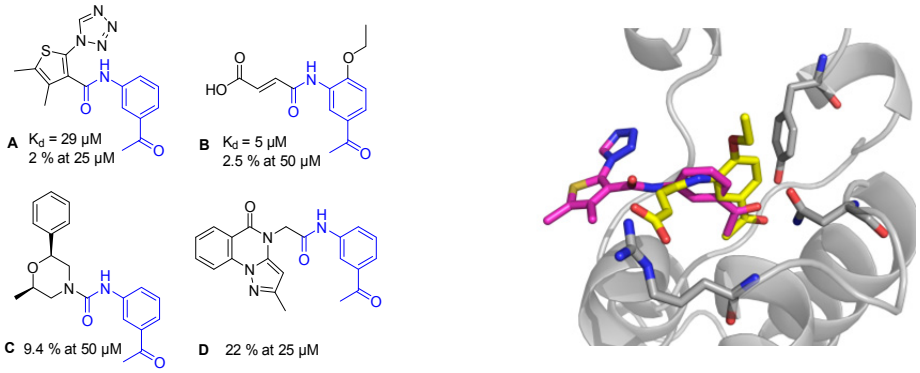


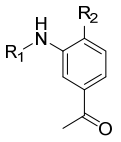

**Figure 23.** Left: Commercially available compound **A** identified in the high-throughput docking campaign and its available analogues **B**-**D**. Equilibrium dissociation constant ($K_d$) values and/or the percentage of residual binding to a known ligand with respect to the DMSO control determined in the competition binding assay (BROMOScan at DiscoveRx)[144-145] are shown. The common structural features of **A**-**D** are shown in blue. Right: Docked poses of **A** (carbon atoms in magenta) and **B** (carbon atoms in yellow) into the CREBBP bromodomain (gray). The side chains of the conserved Tyr1125 and Asn1168 residues, together with Arg1173 are shown as sticks.

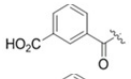
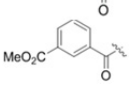
The good affinity of compound **B** for CREBBP ($K_d$ of 5 μM determined by a competition binding assay[148, 149]) highlighted the importance of the salt bridge formed between the molecule's carboxylic acid and the Arg1173 residue (Figure 23, right). Due to its potency, ligand efficiency (0.36 kcal/mol per heavy atom) and chemically novel blueprint within the CREBBP bromodomain inhibitors' pool[131] we decided to focus our optimization campaign on

compound **B**. Our strategy was to preserve its acylbenzene moiety ("head group") since the carbonyl group is involved in a hydrogen bond with the conserved Asn1168 in the docked pose.

Several modifications were designed in order to generate both potent and selective CREBBP inhibitors. First, we decided to keep the fumaric acid moiety and modulate the interactions of the substituents in *ortho-* position relative to the amide group with the hydrophobic residues located on top of the binding site, Ile1122 and Leu1120. Replacement of the ethoxy group originally present in **B** with OMe, Me or F substituents did not have a strong effect on affinity (Table 2, compounds **40-43**). We then turned our attention towards the side chain of the molecule while preserving the carboxylic acid moiety that interacts with Arg1173. Remarkably, the substitution of the fumaric acid by an isophthalic group in **45** (which might also provide a higher metabolic stability) resulted in a 6-fold improvement in binding affinity representing the first nanomolar CREBBP inhibitor of our derivatization campaign. We hypothesized that such improvement in affinity could be due to a more favorable interaction with the Arg1173 residue and the so-called LPF shelf together with a less unfavorable entropic penalty compared to the slightly more flexible fumaric acid derivative. The methyl isophthalate derivative **46** showed similar potency as **45** indicating that its ester is involved in favorable polar interactions with Arg1173.

**Table 2.** First approach towards the optimization of compound **B**.




| Cmpd | $R_1$ | $R_2$ | $\Delta$Tm(°C)[a] | $K_d$ ($\mu$M)[b] | $IC_{50}$ ($\mu$M)[c] |
|---|---|---|---|---|---|
| **B** | Fumaric acid | OEt | 1.8 | 5.00 | – |
| **40** | Fumaric acid | OMe | 1.1 | – | 60 |
| **41** | Fumaric acid | Me | 0.2 | – | >100 |
| **42** | Fumaric acid | F | -0.1 | – | >100 |
| **43** | Fumaric acid | H | 0.1 | – | 30 |
| **44** | 1,2-Cyclopropyl carboxylic acid | OEt | 1.1 | – | >100 |
| **45** |  | OEt | 3.8 | 0.77 | 8.7 |
| **46** |  | OEt | 2.8 | 0.59 | 5.0 |
| **47** |  | OEt | 1.9 | – | 64 |





[a] Median value of the shift in the melting temperature (number of measurements > 10). SEM values did not exceed 0.2 °C a part from the SEM value for compound 7 (0.5°C). [b] $K_d$ values were determined by a competition binding assay[148, 149] in duplicates. [c] $IC_{50}$ values were determined at BPS Bioscience by means of a TR-FRET assay in duplicates.

The crystal structure of compound **45** in complex with the CREBBP bromodomain (solved at a 1.7 Å resolution, Figure 24) revealed an overall binding mode similar to the docked pose of the fumaric acid derivative **B** obtained *in silico* (Figure 23, right). Compound **45** binds in the acetyl-lysine pocket,[36] with its acyl substituent involved in a hydrogen bond with the side chain of the conserved Asn1168 and a water-bridged hydrogen bond with the Tyr1125 side chain hydroxyl. The acetyl benzene head, together with its ethoxy substituent, presents high shape complementarity with the binding site. The NH group of the amide linker forms a hydrogen bond with a water molecule whose oxygen atom is located at a distance of 3.0 Å from the carbonyl oxygen of Pro1110 (in the LPF shelf): the geometry for a water-bridged hydrogen bond, however, is not ideal as the water oxygen is not in the plane of the Pro110 carbonyl. In contrast, the carbonyl group of the amide linker is involved in a favorable water-bridged hydrogen bond with the guanidinium of the Arg1173 side chain. Crucially, the aromaticity of the benzoic acid provides the optimal orientation for the carboxylic acid of compound **45** to form a very favorable polar interaction with the Arg1173 guanidinium of the CREBBP bromodomain. Previously, electrostatic interactions between the sulfonate group of ischemin and this arginine residue were observed.[124] Moreover, Conway and coworkers already reported a relatively weak ligand of CREBBP and BET bromodomains bearing a carboxylic acid that, even if not in direct contact with Arg1173, could form weak electrostatic interactions with this positively charged residue, resulting in a preferential binding towards CREBBP.[68]

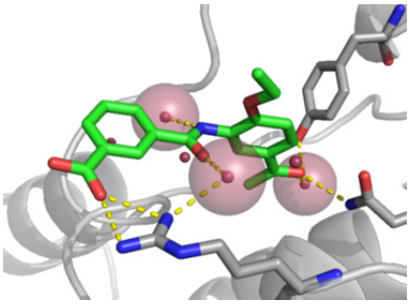


**Figure 24.** Crystal structure of CREBBP (gray) in complex with compound **45** (green) (PDB code 4TQN). The conserved Tyr1125 and Asn1168 residues, together with Arg1173 are shown as sticks.

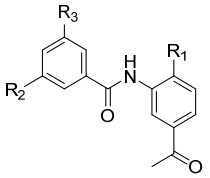
Finite-difference Poisson calculations were performed with CHARMM[146-147] using the crystal structure of the CREBBP bromodomain in complex with compound **45** (PDB code 4TQN) to evaluate the electrostatic contribution of the polar interaction to the free energy of binding of compounds **45** and **46**. The finite-difference Poisson calculation takes into account solvent screening effects which are significant because the Arg1173 side chain is partially exposed to the solvent (Figure 24). Despite the partial solvent accessibility, the electrostatic interaction between the carboxyl group of compound **45** and the guanidinium group of Arg1173 contributes -13.6 kcal/mol, which is about half of the total electrostatic interaction energy (-26.9 kcal/mol) between compound **45** and the CREBBP bromodomain.[148] Analysis of the

individual contributions to the binding free energy of compounds **45** and its methyl ester derivative **46** shows that the former has a slightly more favorable electrostatic contribution ($\Delta G^{45}_{electr} - \Delta G^{46}_{electr}$ = -0.8 kcal/mol) whereas the latter has a more favorable van der Waals contribution ($\Delta G^{45}_{vdW} - \Delta G^{46}_{vdW}$ = 1.4 kcal/mol) so that the difference in the calculated total binding free energy is close to zero in line with the very similar dissociation constants of compounds **45** and **46** measured experimentally (Table 3).

With the high resolution crystal structure of the first nM ligand **45** in hand we further explored the chemical space around this scaffold by introducing modifications at three different positions as illustrated in Table 3. Thermal shift values similar to the one of compound **45** were observed for ligands bearing larger hydrophobic substituents within the acyl benzene head (**48**-**51**). More importantly, the replacement of the carboxylic acid by a tetrazole as isoster (**54**) maintained the affinity, whereas sulfonamides (**55**, **56**) could not improve the large affinity gain achieved by the salt bridge formed in the presence of carboxylic acid in compound **45**. We then decided to enlarge our inhibitors towards the ZA channel by introducing both heteroaromatic (**57**, **58**) as well as linear substituents (**59**-**61**) bearing hydrogen bond acceptors or donors in $R_3$ that could interact with the amino acid residues located in the ZA loop. These modifications led to the most potent inhibitors of our optimization campaign, compounds **58**, **60** and **61**, with $K_d$ values of 170 nM, 540 nM and 400 nM, respectively (Table 4). The similar thermal shift values measured with the CREBBP bromodomain and its paralogue, EP300 (Table 3), are consistent with the fact that identical residues are present in the acetyl-lysine binding site of both proteins.

Compounds **62** and **63** were synthesized (Figure 25, top) to explore the effect of the amide direction (**62**) and the presence of an N-methyl carboxamide as a surrogate of the acetyl moiety (**63**).[131] In both cases a decrease in affinity for CREBBP was observed, confirming the excellent shape complementarity between the acyl benzene moiety and the acetyl-lysine binding site of the CREBBP bromodomain and its paralogue EP300.

**Table 3.** Optimization of compound **45**.




| Cmpd | $R_1$ | $R_2$ | $R_3$ | $\Delta T_m$ (°C)[a] | | $IC_{50}$ |
|---|---|---|---|---|---|---|
| | | | | CBP | EP300 | (μM) [b] |
| **48** | OPr | $CO_2H$ | H | 3.8 | 3.4 | 7.4 |
| **49** | O$^i$Bu | $CO_2H$ | H | 3.5 | 3.0 | NA |
| **50** | $OCH_2Cyc$ | $CO_2H$ | H | 3.4 | 3.6 | NA |
| **51** | OBn | $CO_2H$ | H | 3.3 | 2.6 | NA |
| **52** | Morph | $CO_2H$ | H | 1.4 | 1.3 | NA |
| **53** | Cyc | $CO_2H$ | H | 3.1 | 3.1 | 15 |
| **54** | OEt | 1-Tetrazole | H | 3.3 | 2.6 | 7.5 |
| **55** | OEt | $PhNHSO_2$ | H | 0.0 | 0.0 | >10 |
| **56** | OEt | $PhSO_2NH$ | H | 0.5 | 0.4 | NA |
| **57** | OEt | $CO_2H$ | 4-pyridyl | 4.1 | 3.9 | 3 |
| **58** | OEt | $CO_2H$ | 3-furyl | 5.2 | 5.9 | 1[c] |
| **59** | OEt | $CO_2H$ | $CH_2OH$ | 3.8 | 3.7 | 3 |
| **60** | OEt | $CO_2H$ | $CH_2OMe$ | 5.1 | 4.6 | 2[c] |
| **61** | OEt | $CO_2H$ | $CH_2OTHP$ | 6.0 | 5.9 | 1[c] |

[a] Median value of the shift in the melting temperature (number of measurements > 4). SEM values did not exceed 0.5 °C. [b] $IC_{50}$ values for the CREBBP bromodomain were determined by means of a TR-FRET assay in duplicates ("NA": an $IC_{50}$ value could not be obtained). [c] $K_d$ values for **58**, **59** and **60** were determined by a competition binding assay[148, 149] in duplicates with values of 0.17, 0.54 and 0.40 μM, respectively. Cyc: cyclopropyl; Morph: N-morpholyl; Bn: benzyl; THP: tetrahydropyrane.

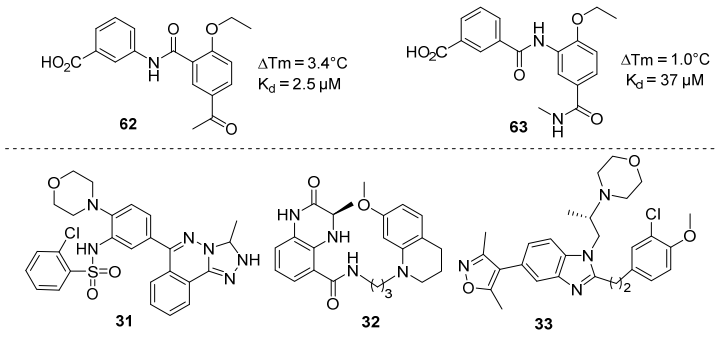



**Figure 25.** Top: Compounds **62** and **63**. Bottom: Potent ligands of the CREBBP bromodomain previously reported.[99, 125-126]

Remarkable target selectivity for CREBBP and EP300 over the other five human bromodomains (each belonging to a different subfamily) was found for the nanomolar ligands **45**, **58**, **60**, and **61** according to thermal shift (see experimental section, Figures 28-30). Interestingly, the most potent inhibitor **58**, showed only 1.7 °C thermal shift against BRD4(1),

which is the most promiscuous bromodomain.[99, 125] The selectivity was further investigated by isothermal titration calorimetry (ITC)[149] and a competition binding assay (Table 4).[148, 149] The BRD4(1)-$K_d$/CREBBP-$K_d$ ratio determined by competition binding yields selectivity factors of >65, 59, and 48 for compounds **45**, **58**, and **60**, respectively. Such selectivity towards the CREBBP bromodomain confirms the importance of the salt bridge with the positively charged Arg1173 for the design of selective CREBBP ligands. Importantly, the high ligand efficiency (LE)[150-151] was preserved throughout the optimization campaign. Moreover, the LE values (0.28 to 0.35 kcal/mol per heavy atom) and the ligand-lipophilic efficiency (LLE) values (close to 4.0) of our nM ligands **45**, **58**, **60**, and **61**, compare positively with those of previously disclosed ligands of the CREBBP bromodomain **31-33** (Figure 25, bottom).[99, 125-126]

**Table 4.** Activity and selectivity of the most potent acyl benzene derivatives.

| Cmpd | LE[a] | LLE | $K_d$ (μM) ITC | $K_d$ (μM) competition binding assay[148, 149] | | | | | $\Delta T_m$ (°C)[b] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | CBP | CBP | BRD4(1) | BRD4(2) | BRD2(1,2) | S[c] | CBP | EP300 | BRD4(1) |
| **45** | 0.35 | 3.7 | 2.0 | 0.77 | >50 | >50 | >50 | >65 | 3.8 | 3.4 | 0.4 |
| **58** | 0.32 | 3.2 | 0.3 | 0.17 | 10 | 36 | 36 | 59 | 5.2 | 5.9 | 1.5 |
| **60** | 0.32 | 4.0 | 0.8 | 0.54 | 26 | >50 | >50 | 48 | 5.1 | 4.6 | 1.3 |
| **61** | 0.28 | 3.8 | – | 0.40 | - | - | - | - | 6.0 | 5.9 | 1.7 |

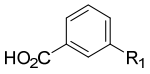
[a] LE = ligand efficiency, calculated as ($\Delta G$/number of heavy atoms), is reported in kcal/mol per heavy atom; LLE = lipophilic ligand efficiency (calculated as $pK_d$-clogP),[150-151] clogP was calculated using ChemDraw. [b] Median value of the shift in the melting temperature (number of measurements > 9). SEM values did not exceed 0.3 °C. [c] Selectivity (S) between the CREBBP and BRD4(1) bromodomains determined by the ratio of $K_d$ values obtained via the competition binding assay.[148, 149]

Screening of our most potent CREBBP inhibitors against a panel of ten cancer cell lines showed lack of overall toxicity and significant growth inhibition for the leukaemia cell lines MOLM-13, ML2 and HL-60 specifically for compounds **45**, **48**, **61** and **62** (see experimental section, Tables 9 and 10). The methyl esters derived from compounds **45**, **48-50**, **58** and **59-62** were also tested showing an equivalent or higher cell growth inhibitory activity compared to the corresponding acid counterparts, probably due to their better cell permeability. In fact, the methyl ester of **61** presented toxicity values in the low micromolar range, with $GI_{50}$ values of 14.4 and 5.3 μM for ML2 and MOLM-13, respectively, which could be linked to the involvement of CREBBP in some subtypes of acute myeloid leukaemia.[152-154]

In an effort to expand the chemical diversity of the thus far limited portfolio of selective CREBBP/EP300 ligands,[68, 99, 123-125] we decided to replace the acetyl benzene head by different chemical motifs while keeping the benzoic acid substitution that proved to be crucial in our optimization campaign. To achieve this goal, a library of about 400 heteroaromatics was docked with SEED[140-141] to the crystal structure of CREBBP (PDB code 3P1C) and BRD4(1)

(PDB code 3MXF). The top fragments sorted according to predicted binding energy to CREBBP were visually inspected while overlaying their docked pose to the X-ray crystal structure of compound **45**. Six fragments were selected based on their binding energy to CREBBP and selectivity against BRD4(1) as predicted by SEED[140-141] (Table 6 in the experimental section), their orientation relative to the benzoic acid moiety of compound **45**, their novelty within the bromodomain ligand pool, and their synthetic accessibility. Hybridization of these fragments with benzoic acid yielded compounds **64-66**, which proved to be low micromolar ligands of the CREBBP bromodomain in the competition binding assay (Table 5). Moreover, the selectivity values of the new CREBBP ligands are consistent with the predicted binding energies calculated *in silico*.[155]

**Table 5.** Benzoic acid-templated ligands containing novel acetylated lysine mimics.




| Cmpd | LE | $K_d$ (µM)[a] | | $\Delta T_m$(°C)[b] | | $\Delta E_{calculated}$[c] | |
|---|---|---|---|---|---|---|---|
| | CBP | CBP | BRD4(1) | CBP | BRD4(1) | CBP | BRD4(1) |
| **64** | 0.30 | 17.0 | 6.9 | 1.2 | 2.6 | -19.9 | -18.3 |
| **65** | 0.37 | 4.1 | 33 | 1.9 | 1.0 | -19.4 | -14.5 |
| **66** | 0.36 | 1.1 | 41 | 4.4 | 1.5 | -19.9 | -16.5 |

[a] $K_d$ values were determined by a competition binding assay[148, 149] in duplicates. A $K_d$ value of 1.6 µM for compound 30 was measured by ITC (Figure 31 in the experimental section). [b] Median value of the shift in the melting temperature (number of measurements > 12). SEM values did not exceed 0.3 °C. [c] SEED total energy[140-141] in kcal/mol for the compounds without amide linker. The compounds are involved in a hydrogen bond with the conserved Asn1168 in the binding modes obtained by docking with SEED.[140-141]

In summary, we have discovered *in silico*, validated by X-ray crystallography, and optimized by chemical synthesis a series of nM potent and selective acyl benzene ligands of the CREBBP bromodomain. Fragment-based, high-throughput docking was employed for the identification of novel scaffolds whose affinity was enhanced in a straightforward manner by incorporating interactions within the ZA channel. The direct polar interaction between the benzoic acid moiety of these ligands (e.g. compound **45**, Figure 24, PDB code 4TQN) and the Arg1173 guanidinium is unprecedented and can be exploited for the design of selective ligands of the CREBBP bromodomain. In a second step of our *in silico* design campaign, we hybridized the benzoic acid moiety with fragments docked close to the conserved Asn1168 and prioritized according to predicted energy of binding to the CREBBP bromodomain, obtaining three novel acetylated lysine mimics. The specificity and easy synthetic availability of our compounds will be useful to unravel the role of CREBBP in several types of solid tumours and haematological malignancies.

## 3.3 Experimental section

### 3.3.1 Fragment-based high-throughput docking

The ALTA (anchor-based library tailoring) procedure was used for fragment-based high-throughput docking (Figure 26). First the program DAIM[139] was employed to decompose the nearly two million compounds in the ZINC leads-now library (version of Oct. 2012). The conserved Asn1168 of the CREBBP bromodomain (numbering from PDB structure 3SVH) is involved in binding the natural ligand acetyl-lysine. To target Asn1168 fragments with at least one hydrogen bond donor or acceptor were kept. Further removal of fragments without a ring resulted in a library of 106,534 fragments.

The crystal structures of the CREBBP bromodomain in complex with the natural ligand acetyl-lysine (PDB code 3P1C) and the complex with the synthetic inhibitor 3,5-dimethylisoxazole (3SVH) show different orientation of the side chain of Val1174, the so-called gatekeeper residue.[156] In addition, the side chain of Arg1173 points towards the solvent and seems very mobile. In the 3P1C structure, Arg1173 is located just above the Leu-Pro-Phe segment (usually called WPF-shelf as Trp is the first residue in this triad in most bromodomains) and acts as a potential hydrogen donor around that region.[129] However, in 3SVH the Arg1173 side chain points in a direction away from the Leu-Pro-Phe segment. Given these differences, it was decided to use both structures for docking.

The library of 106,534 fragments was docked by the program SEED,[141] which calculates the electrostatic energy in a continuum approximation using an efficient numerical evaluation of the Born radii. The Born radii are a measure of the amount of solute (i.e., low-dielectric volume) around each atom of the protein or fragment. They are employed in the generalized Born formula for the calculation of the electrostatic contribution to the intermolecular interaction and desolvation of the ligand.[142, 157]

After fragment docking by the program SEED two additional filters were applied to reduce the number of fragments to less than 5,000 for each of the two CREBBP inhibitors: SEED total energy efficiency more favorable than -0.125 kcal $mol^{-1}$ $g^{-1}$ and hydrogen bonding penalty less or equal 1.[158] With these filters, 91% (docked to 3SVH) and 82% (docked to 3P1C) of the fragments were involved in hydrogen bond(s) with Asn1168 while these percentages were 33% (3SVH) and 41% (3P1C) for the entire dataset.

The top (about 4,000) fragments were used to retrieve the molecules in the ZINC library which they originated from (called parent compounds hereafter). There were 808,588 (docking in 3SVH) and 754,411 (3P1C) parent compounds in total and for each of them at most 15 poses were generated by AutoDock Vina.[143] All the poses of the parent compounds were then

minimized, filtered and rescored. In the end, the 1,000 top-ranked poses according to the SEED total energy were clustered based on the functional groups interacting with the conserved Asn1168. This clustering yielded only 20 different functional groups. Finally, explicit solvent molecular dynamics simulations (started from the docked poses) were used to further reduce the number of compounds to 17 representative molecules which were purchased for validation *in vitro*.

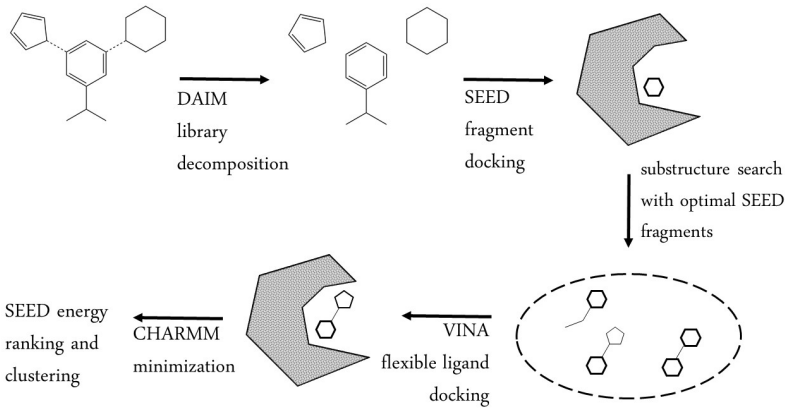



**Figure 26.** Application of the ALTA (anchor-based library tailoring) procedure[136, 138] to the CREBBP bromodomain. The program DAIM automatically decomposes molecules into fragments by cutting at rotatable bonds.[139] CHARMM atom types and non-bonding parameters used in SEED[141-142, 157] were generated by MATCH.[159] The program for fragment docking (called SEED[141-142, 157]) requires about 5 seconds per fragment on a single core of an i7 CPU at 2.8 GHz. CHARMM minimization of the fragment with rigid CREBBP took about 2 seconds per fragment.

### 3.3.2 Finite-difference Poisson calculations

The electrostatic contribution to the binding free energy was evaluated by numerical solution of the Poisson equation using the finite-difference method as implemented in the *PBEQ* module[147] of the program CHARMM.[146] The solute/solvent dielectric discontinuity surface was delimited by the molecular surface spanned by the surface of a rolling probe of 1.4 Å. The dielectric constant of the solute and the solvent were set to 2.0 and 78.5, respectively. The ionic strength was set to zero, and the temperature was set to 300 K. The size of the initial grid was determined by considering a layer of at least 20 Å around the solute. The partial charges of the solute were distributed on the grid points by the trilinear interpolation algorithm. First the linearized Poisson equation was solved on a grid of 1.0 Å spacing, which was followed by a focused calculation with a grid encompassing all of the solute and a grid spacing of 0.3 Å. For both calculations an iterative procedure (successive over-relaxation) was used. All calculations were carried out independently on the crystal structure (PDB code 4TQN) and the minimized crystal structure. The latter was obtained by conjugate gradient minimization with the CHARMM program (version 38b1), and the CHARMM param36 force

field for CREBBP and the CHARMM general force field (which is compatible with CHARMM param36) for compound **45**.

### 3.3.3 Fragment docking

Fragment docking to the CREBBP and BRD4(1) bromodomain was carried out with the in-house developed program SEED.

**Table 6.** Selected fragments to be coupled to the benzoic acid side chain and their binding energy values calculated by SEED.[140-141]

| | CREBBP | | | BRD4(1) | | |
|---|---|---|---|---|---|---|
| Fragment | van der Waals | electrostatic | total | van der Waals | electrostatic | total |
| **I** | -14.0 | -2.3 | -16.3 | -14.6 | 0.5 | -14.0 |
| **II** | -18.5 | -1.4 | -19.9 | -17.5 | -0.8 | -18.3 |
| **III** | -17.5 | -1.9 | -19.4 | -16.6 | 2.1 | -14.5 |
| **IV** | -15.3 | -4.6 | -19.9 | -17.7 | 1.1 | -16.5 |
| **V** | -16.7 | -1.4 | -18.1 | -15.9 | 1.6 | -14.4 |
| **VI** | -15.3 | -3.1 | -18.4 | -15.8 | 1.8 | -14.1 |
| **VII** | -15.6 | -4.0 | -19.6 | -15.4 | -0.6 | -16.1 |

The CHARMM36[160] and CGenFF[161] force fields were used for the protein and fragments, respectively. The electrostatic energy is the sum of bromodomain desolvation, fragment desolvation, and intermolecular interaction screened by the solvent which is treated implicitly by the generalized Born approximation.[162]. All values are in kcal/mol.

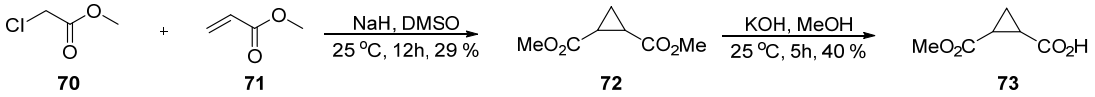
### 3.3.4 Synthetic methods

All reactions, unless otherwise stated, were carried out under inert gas atmosphere using standard Schlenk-techniques. All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 $F_{254}$. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV 300, AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in ppm. The spectra are calibrated to the residual $^{1}H$ and $^{13}C$ signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Mettler Toledo MP70 melting point instrument. High-resolution electrospray ionization mass spectrometry was performed on a

Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. Ten spectra were acquired. A mass accuracy ≤2 ppm was obtained in the peak matching acquisition mode by using a solution containing 2 µL PEG200, 2 µL PPG450, and 1.5 mg NaOAc (all obtained from Sigma-Aldrich, Buchs, Switzerland) dissolved in 100 mL MeOH (HPLC Supra grade, Scharlau, E-Barcelona) as internal standard.

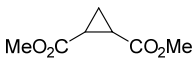
The following compounds were prepared according to previously reported procedures: **72**,[163] **73**,[164] **75**,[165] **80**,[166] **81**,[167] **82**,[167] **84**,[168] **86**,[169] **88**,[170] **92**[171] and **93**.[172]

## I. Synthesis of non-commercially available carboxylic acids (37-57)



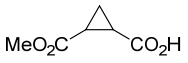

**70** **71** **72** **73**

### Dimethyl cyclopropane-1,2-dicarboxylate (72)[163]

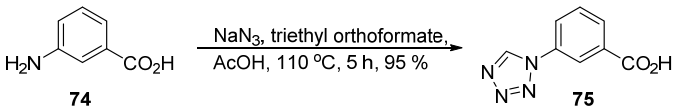



Colourless oil; Yield: 29 %; $^1$H NMR (400 MHz, $CDCl_3$): δ = 3.70 (s, 6H), 2.22 – 2.12 (m, 2H), 1.48 – 1.40 (m, 2H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 172.2, 52.1, 22.2, 15.3; IR (neat): ῦ = 2955, 1724, 1437, 1399, 1332, 1270, 1197, 1169, 1026, 904, 839, 753, 661, 444, 419, 407 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_7H_{10}NaO_4^+$, 181.1; found, 180.9.

### 2-(Methoxycarbonyl)cyclopropane-1-carboxylic acid (73)[164]

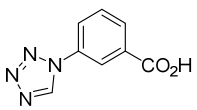



Colourless oil; Yield: 40 %; $^1$H NMR (400 MHz, $CDCl_3$): δ = 3.72 (s, 3H), 2.27 – 2.22 (m, 1H), 2.21 – 2.16 (m, 1H), 1.50 (tdd, *J* = 9.9, 6.0, 4.0 Hz, 2H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 177.5, 171.8, 52.3, 22.7, 22.0, 15.8; IR (neat): ῦ = 2958, 1704, 1438, 1309, 1176, 913, 743, 415 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_6H_7O_4^-$, [M-H]$^-$, 143.0; found, 142.9.

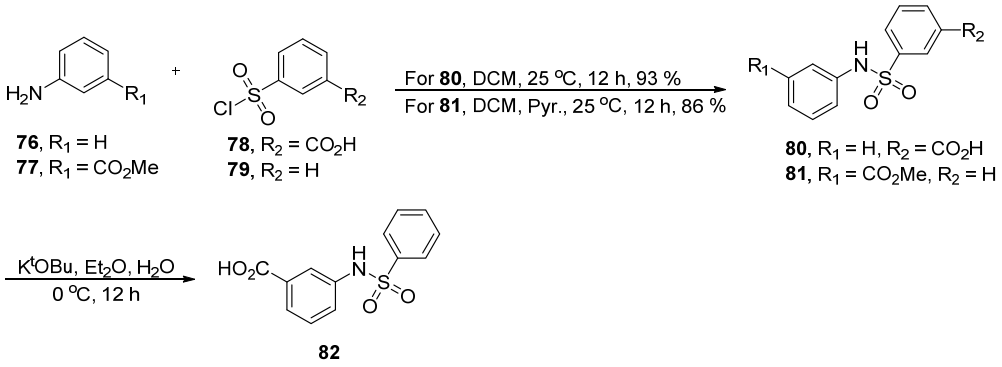



**74** **75**

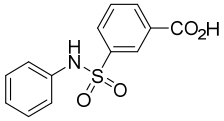
### 3-(1*H*-Tetrazol-1-yl)benzoic acid (75)[165]




White solid; Yield: 95 %; mp 174-176 °C; $^1$H NMR (300 MHz, DMSO-$d_6$): δ = 13.48 (br, 1H), 10.21 (s, 1H), 8.42 (s, 1H), 8.18 (d, *J* = 8.0 Hz, 1H), 8.10 (d, *J* = 7.8 Hz, 1H), 7.78 (t, *J* = 7.9 Hz, 1H); $^{13}$C NMR (75 MHz, DMSO-$d_6$): δ = 166.1, 142.4, 134.0, 132.7, 130.6, 130.2, 125.2, 121.6; IR (neat): ῦ = 3454, 3397, 3137, 2898, 2788, 2621, 2518, 1704, 1636, 1592, 1500, 1479, 1453, 1289, 1277, 1217, 1195, 1103, 1063, 900, 816, 758, 707 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_8H_5N_4O_2^-$, [M-H]$^-$, 189.0; found, 188.8.

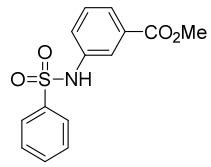



### 3-(*N*-Phenylsulfamoyl)benzoic acid (80)[166]



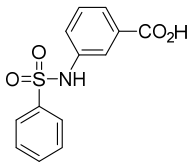
White solid; Yield: 93 %; mp 193-195 °C; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 10.36 (s, 1H), 8.29 (t, $J$ = 1.7 Hz, 1H), 8.12 (d, $J$ = 7.8 Hz, 1H), 7.94 (d, $J$ = 7.9 Hz, 1H), 7.67 (t, $J$ = 7.8 Hz, 1H), 7.23 (t, $J$ = 7.9 Hz, 2H), 7.11 – 7.01 (m, 3H); $^{13}$C NMR (101 MHz, DMSO-$d_6$): δ = 165.9, 140.0, 137.3, 133.3, 131.8, 130.6, 129.9, 129.2, 127.3, 124.5, 120.5; IR (neat): ῡ = 3246, 2980, 2868, 1686, 1598, 1493, 1444, 1408, 1339, 1294, 1218, 1207, 1173, 1159, 1135, 1075, 1025, 928, 899, 854, 823, 749, 735 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{13}H_{10}NO_4S^-$, [M-H]$^-$, 276.0; found, 275.9.

### Methyl 3-(phenylsulfonamido)benzoate (81)[167]

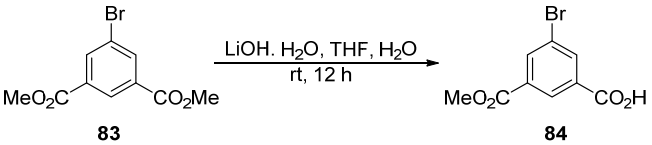


White solid; Yield: 86 %; mp 161-163 °C; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 10.55 (s, 1H), 7.76 (d, $J$ = 7.4 Hz, 2H), 7.70 (s, 1H), 7.61 (d, $J$ = 7.0 Hz, 2H), 7.57 – 7.54 (m, 2H), 7.42 – 7.34 (m, 2H), 3.81 (s, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ = 165.6, 139.2, 138.2, 133.1, 130.5, 129.8, 129.4, 126.6, 124.7, 124.4, 120.3, 52.3; IR (neat): ῡ = 3412, 3227, 3070, 1701, 1605, 1591, 1474, 1446, 1438, 1401, 1337, 1296, 1216, 1175, 1156, 1122, 1114, 1088, 1000, 984, 948, 901, 853, 752, 714 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{13}NNaO_4S^+$, 314.1; found, 314.0.

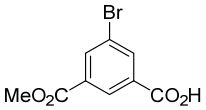
### 3-(Phenylsulfonamido)benzoic acid (82)[167]



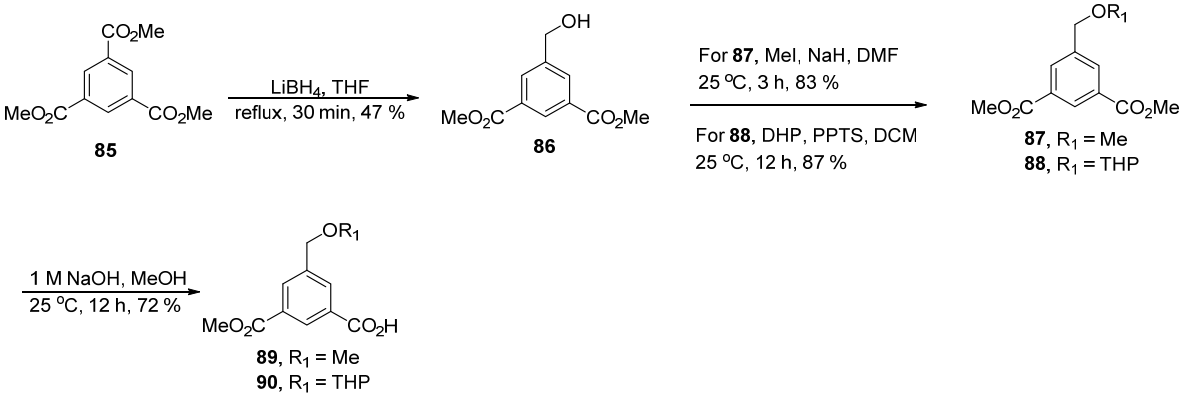
Acid **82** was prepared using the reported procedure and used without further purification. The presence of the acid was confirmed by IR and MS. IR (neat): ῡ = 3250, 3064, 2963, 2825, 1682, 1587, 1447, 1424, 1405, 1333, 1297, 1263, 1158, 1089, 948, 906, 882, 823, 762, 749, 722 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{13}H_{10}NO_4S^-$, [M-H]$^-$, 276.0; found, 276.0.



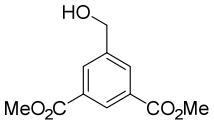
**3-Bromo-5-(methoxycarbonyl)benzoic acid (84)**[168]



Acid **84** was prepared using the reported procedure and used without further purification. The presence of the acid was confirmed by IR and MS. IR (neat): $\tilde{\upsilon}$ = 3079, 2859, 2647, 2522, 1684, 1600, 1574, 1540, 1448, 1391, 1267, 1199, 1147, 905, 797, 754, 727, 712 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_9H_6BrO_4^-$, $[M-H]^-$, 256.9; found, 256.8.

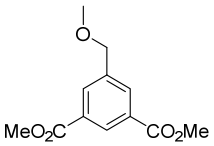



**Dimethyl 5-(hydroxymethyl)isophthalate (86)**[169]



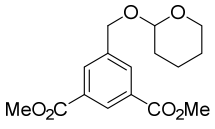
White solid; Yield: 47 %; mp 106-108 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 8.60 (s, 1H), 8.31 – 8.15 (m, 2H), 4.81 (s, 2H), 3.95 (s, 6H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 166.2, 141.9, 132.0, 130.9, 129.8, 64.2, 52.4; IR (neat): $\tilde{\upsilon}$ = 3490, 2954, 1719, 1603, 1437, 1247, 1183, 1107, 1002, 969, 921, 888, 788, 749, 719 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{11}H_{12}NaO_5^+$, 247.1; found, 246.9.

**Dimethyl 5-(methoxymethyl)isophthalate (87)**



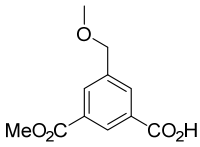
To a solution of dimethyl 5-(hydroxymethyl)isophthalate (**86**, 150 mg, 0.669 mmol) in DMF (1.50 mL) NaH (53.5 mg, 1.34 mmol) was added and stirred at 25 °C for 30 minutes. Methyl iodide (83.3 µL, 1.34 mmol) was added and the reaction mixture was stirred for 3 hours at 25 °C. The reaction was quenched by addition of 1 M HCl solution and extracted with EtOAc three times. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was then purified by flash column chromatography (hexane: EtOAc, 3:1) affording the desired product in pure form as a white solid (133 mg, 0.558 mmol, 83 % yield). mp 83-85 °C; $^1$H NMR (500 MHz, $CDCl_3$): δ = 8.60 (s, 1H), 8.28 – 8.13 (m, 2H), 4.54 (s, 2H), 3.95 (s, 6H), 3.42 (s, 3H); $^{13}$C NMR (126 MHz, $CDCl_3$): δ = 166.2, 139.4, 132.8 130.8, 130.0, 73.6, 58.5, 52.4; IR (neat): $\tilde{\upsilon}$ = 3490, 2954, 1719, 1603, 1437, 1247, 1183, 1107, 1002, 969, 921, 888, 788, 749, 719 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{12}H_{14}NaO_5^+$, 261.1; found, 261.0.

**Dimethyl 5-((tetrahydro-2*H*-pyran-2-yl)methyl)isophthalate (88)**[170]



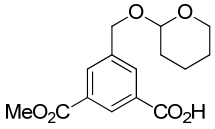
White solid; Yield: 87 %; mp 45-48 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 8.59 (t, *J* = 1.6 Hz, 1H), 8.25 – 8.19 (m, 2H), 4.85 (dd, *J* = 12.4, 0.4 Hz, 1H), 4.73 (t, *J* = 3.5 Hz, 1H), 4.57 (d, *J* = 12.4 Hz, 1H), 3.94 (s, 6H), 3.93 – 3.85 (m, 1H), 3.60 – 3.51 (m, 1H), 1.93 – 1.81 (m, 1H), 1.81 – 1.51 (m, 5H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 166.2, 139.5, 132.9, 130.7, 129.8, 98.2, 67.9, 62.2, 52.3, 30.4, 25.4, 19.2; IR (neat): ῡ = 2952, 2927, 2882, 2846, 1719, 1603, 1453, 1434, 1331, 1318, 1242, 1214, 1203, 1108, 1055, 1015, 965, 911, 869, 811, 751, 720, 712 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{16}H_{20}NaO_6^+$, 331.1; found, 331.1.

**3-(Methoxycarbonyl)-5-(methoxymethyl)benzoic acid (89)**[173]

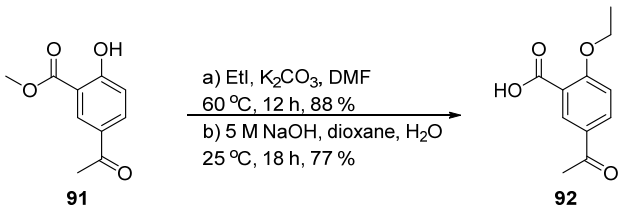


Acid **89** was prepared using the reported procedure and used without further purification. The presence of the acid was confirmed by IR and MS. IR (neat): ῡ = 2918, 2861, 2824, 2626, 1694, 1605, 1458, 1440, 1417, 1314, 1263, 1209, 1115, 1103, 1003, 968, 922, 751, 704 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{11}H_{11}O_5^-$, $[M-H]^-$, 223.1; found, 222.9.

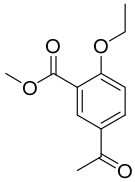
**3-(Methoxycarbonyl)-5-(((tetrahydro-2*H*-pyran-2-yl)oxy)methyl)benzoic acid (90)**



To a solution of dimethyl 5-((tetrahydro-2*H*-pyran-2-yl)methyl)isophthalate (**88**, 150 mg, 0.486 mmol) 1M NaOH solution (0.486 mL, 0.486 mmol) was added and stirred at 25 °C for 12 hours. The pH of the reaction mixture was then adjusted to 3-4 by adding 10 % citric acid solution and extracted with EtOAc three times and brine. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was then purified by flash column chromatography (5 % MeOH in DCM) affording the desired product in pure form as a white solid (102 mg, 0.347 mmol, 72 %). mp 87-93 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 10.92 (br, 1H), 10.92 (s, 1H), 8.30 (s, 1H), 8.28 (s, 1H), 4.88 (d, *J* = 12.5 Hz, 1H), 4.76 (t, *J* = 3.3 Hz, 1H), 4.60 (d, *J* = 12.5 Hz, 1H), 3.95 (s, 3H), 3.94 – 3.87 (m, 1H), 3.63 – 3.53 (m, 1H), 1.88 (dt, *J* = 13.0, 4.3 Hz, 1H), 1.83 – 1.49 (m, 5H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 170.7, 166.1, 139.7, 133.7, 133.4, 130.9, 130.5, 130.0, 98.2, 67.9, 62.2, 52.4, 30.4, 25.3, 19.2; IR (neat): ῡ = 2943, 2873, 1727, 1687, 1606, 1458, 1436, 1421, 1354, 1310, 1252, 1202, 1121, 1072, 1034, 1017, 975, 942, 912, 869, 814, 752, 703 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{15}H_{18}NaO_6^+$, 317.0996; found, 317.0994.

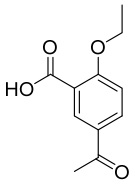


**Methyl 5-acetyl-2ethoxybenzoate (92)**[171]



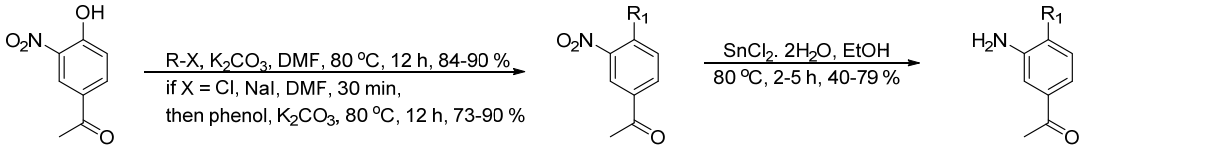
White solid; Yield: 88 %; mp 50-53 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 8.38 (t, $J$ = 2.2 Hz, 1H), 8.08 (dd, $J$ = 8.8, 2.4 Hz, 1H), 7.00 (d, $J$ = 8.8 Hz, 1H), 4.20 (q, $J$ = 7.0 Hz, 2H), 3.91 (s, 3H), 2.57 (d, $J$ = 2.0 Hz, 3H), 1.49 (t, $J$ = 7.0 Hz, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 196.1, 166.0, 162.0, 133.6, 132.6, 129.4, 120.1, 112.6, 64.9, 52.1, 26.3, 14.5; IR (neat): ῡ = 2986, 2947, 1697, 1676, 1597, 1496, 1468, 1439, 1352, 1266, 1249, 1233, 1164, 1110, 1101, 1074, 1036, 958, 929, 859, 841, 816, 786, 718 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{12}H_{14}NaO_4^+$, 245.1; found, 244.9.

**5-Acetyl-2-ethoxybenzoic acid (93)**[172]



Pale yellow solid; Yield: 77 %; mp 118-121 °C; $^1$H NMR (400 MHz, MeOD): δ = 8.39 (d, $J$ = 2.4 Hz, 1H), 8.13 (dd, $J$ = 8.8, 2.4 Hz, 1H), 7.18 (d, $J$ = 8.8 Hz, 1H), 4.24 (q, $J$ = 7.0 Hz, 2H), 2.57 (s, 3H), 1.46 (t, $J$ = 7.0 Hz, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 198.5, 169.0, 163.3, 135.1, 133.4, 130.6, 121.9, 114.0, 66.2, 26.4, 14.8; IR (neat): ῡ = 2989, 2938, 2810, 2615, 1681, 1666, 1598, 1497, 1403, 1354, 1293, 1263, 1224, 1167, 1110, 1076, 1027, 968, 928, 806 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{11}H_{11}NaO_4^+$, 230.1; found, 230.9.

## II. Synthesis of non-commercially available anilines (102-108)




**94**

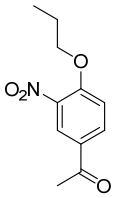
**95**, $R_1$ = propoxy
**96**, $R_1$ = isobutoxy
**97**, $R_1$ = cyclopropylmethoxy
**98**, $R_1$ = benzyloxy
**99**, $R_1$ = methoxy (commercially available)
**100**, $R_1$ = methyl (commercially available)
**101**, $R_1$ = fluoro (commercially available)

**102**, $R_1$ = propoxy
**103**, $R_1$ = isobutoxy
**104**, $R_1$ = cyclopropylmethoxy
**105**, $R_1$ = benzyloxy
**106**, $R_1$ = methoxy
**107**, $R_1$ = methyl
**108**, $R_1$ = fluoro

**General procedure for ether synthesis (95-98)**

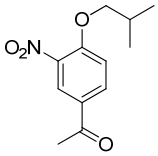
To a solution of 1-(4-hydroxy-3-nitrophenyl)ethan-1-one (**94**, 1.0 eq) in DMF (0.30 M), the alkyl iodide (1.2 eq) and $K_2CO_3$ (4.0 eq) were added. The reaction mixture was stirred at 80 °C for 12 h. A saturated solution of $NH_4Cl$ was added and it was extracted with $Et_2O$ three times. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (2:1, hexane:EtOAc) affording the desired products in pure form. In the case of alkyl chlorides (products **96** and **97**), the alkyl chloride (2.0 eq) was first stirred in DMF (0.30 M) in the presence of NaI (2.0 eq) for 30 min. The phenol (1.0 eq) and $K_2CO_3$ (4.0 eq) were then added and the same procedure as the one described above was followed.

**1-(3-Nitro-4-propoxyphenyl)ethan-1-one (95)**



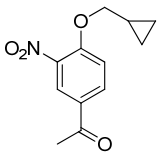
White solid; Yield: 84 %; mp 57-60 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 8.40 (d, *J* = 2.1 Hz, 1H), 8.13 (dd, *J* = 8.8, 2.2 Hz, 1H), 7.12 (d, *J* = 8.8 Hz, 1H), 4.15 (t, *J* = 6.4 Hz, 2H), 2.59 (s, 3H), 2.17 – 1.53 (m, 2H), 1.08 (t, *J* = 7.4 Hz, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 194.8, 155.8, 139.5, 133.8, 129.3, 126.1, 114.0, 71.6, 26.3, 22.2, 10.3; IR (neat): ῡ = 2981, 2972, 2927, 2884, 1666, 1608, 1524, 1499, 1462, 1399, 1356, 1282, 1258, 1156, 1077, 1056, 965, 907, 899, 828, 774 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{11}H_{14}NO_4^+$, 224.0917; found, 224.0918.

**1-(4-Isobutoxy-3-nitrophenyl)ethan-1-one (96)**



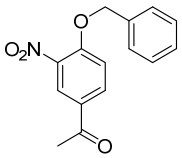
Pale yellow solid; Yield: 90 %; mp 61-63 °C; $^{1}$H NMR (500 MHz, $CDCl_3$): δ = 8.41 (d, *J* = 2.2 Hz, 1H), 8.13 (dd, *J* = 8.8, 2.2 Hz, 1H), 7.11 (d, *J* = 8.8 Hz, 1H), 3.94 (d, *J* = 6.4 Hz, 2H), 2.59 (s, 3H), 2.17 (dp, *J* = 13.3, 6.6 Hz, 1H), 1.07 (s, 3H), 1.06 (s, 3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ = 194.8, 155.9, 139.3, 133.8, 129.2, 126.2, 113.9, 76.1, 28.2, 26.3, 19.0; IR (neat): ῡ = 2966, 2876, 1677, 1609, 1568, 1531, 1470, 1419, 1355, 1273, 1235, 1166, 1066, 1004, 988, 976, 957, 910, 824, 815, 762, 739 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{16}NO_4^+$, 238.1074; found, 238.1077.

**1-(4-(Cyclopropylmethoxy)-3-nitrophenyl)ethan-1-one (97)**



Yellow solid; Yield: 73 %; mp 74-77 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 8.40 (d, *J* = 2.2 Hz, 1H), 8.12 (dd, *J* = 8.8, 2.2 Hz, 1H), 7.11 (d, *J* = 8.8 Hz, 1H), 4.06 (d, *J* = 6.8 Hz, 2H), 2.59 (s, 3H), 1.39 – 1.25 (m, 1H), 0.73 – 0.66 (m, 2H), 0.45 – 0.39 (m, 2H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 194.8, 155.7, 139.6, 133.7, 129.4, 126.1, 114.3, 74.6, 26.3, 9.8, 3.4; IR (neat): ῡ = 3275, 3092, 3011, 1673, 1607, 1566, 1527, 1497, 1412, 1360, 1271, 1171, 1066, 1024, 977, 911, 886, 835, 826, 809 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{14}NO_4^+$, 236.0917; found, 236.0918.

**1-(4-(Benzyloxy)-3-nitrophenyl)ethan-1-one (98)**

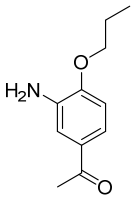


Pale yellow solid; Yield: 90 %; mp 132-135 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 8.42 (d, *J* = 2.2 Hz, 1H), 8.11 (dd, *J* = 8.8, 2.2 Hz, 1H), 7.45-7.32 (m, 5H), 7.18 (d, *J* = 8.8 Hz, 1H), 5.31 (s, 2H), 2.58 (s, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 194.7, 155.1, 139.7, 134.7, 133.7, 129.8, 128.8, 128.5, 126.9, 126.1, 114.6, 71.4, 26.3; IR (neat): ῡ = 2923, 1679, 1611, 1569, 1531, 1493, 1417, 1346, 1267, 1237, 1178, 1066, 1018, 979, 911, 890, 828, 732 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{15}H_{14}NO_4^+$, 272.0917; found, 272.0917.

**General procedure for the reduction of nitro phenyls (102-108)**

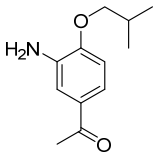
To a mixture of nitrophenyl (1 eq) in EtOH (0.3 M) $SnCl_2 \cdot 2H_2O$ (4 eq) was added. The reaction mixture was heated to 80 °C for 2-5 h, cooled, and concentrated under reduced pressure. The pH was basified to pH 5 by the addition of a 5 M NaOH solution. The resulting precipitate was filtered off and washed with EtOAc. The organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (hex: EtOAc, 3:1-1:1) affording the desired anilines in pure form.

**1-(3-Amino-4-propoxyphenyl)ethan-1-one (102)**



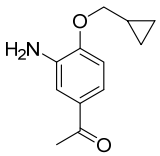
Brown solid; Yield: 40%; mp 53-55 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 7.40 – 7.37 (m, 2H), 6.79 (d, $J$ = 8.1 Hz, 1H), 4.03 (t, $J$ = 6.5 Hz, 2H), 2.52 (s, 3H), 1.98 – 1.78 (m, 2H), 1.07 (t, $J$ = 7.4 Hz, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 197.2, 150.9, 135.8, 130.5, 120.8, 114.3, 110.1, 70.0, 26.3, 22.5, 10.5; IR (neat): $\tilde{\upsilon}$ = 3485, 3369, 2961, 2934, 2877, 1671, 1615, 1580, 1513, 1474, 1439, 1350, 1295, 1253, 1219, 1148, 1063, 1039, 1013, 977, 916, 886, 787, 770 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{11}H_{16}NO_2^+$, 194.1176; found, 194.1179.

**1-(3-Amino-4-isobutoxyphenyl)ethan-1-one (103)**



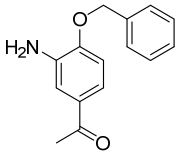
Brown-red oil; Yield: 53 %; $^1$H NMR (400 MHz, $CDCl_3$): δ = 7.37 – 7.34 (m, 3H), 6.77 (d, $J$ = 8.3 Hz, 1H), 3.90 (br, 2H), 3.82 (d, $J$ = 6.5 Hz, 1H), 2.52 (s, 3H), 2.15 (dp, $J$ = 13.3, 6.7 Hz, 1H), 1.06 (s, 3H), 1.05 (s, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 197.3, 150.7, 136.2, 130.5, 120.4, 113.9, 110.0, 74.7, 28.2, 26.2, 19.2; IR (neat): $\tilde{\upsilon}$ = 3472, 3363, 2959, 2928, 2873, 1667, 1613, 1583, 1513, 1470, 1440, 1359, 1296, 1210, 1153, 1064, 1022, 880, 795 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{18}NO_2^+$, 208.1332; found, 208.1334.

**1-(3-Amino-4-(cyclopropylmethoxy)phenyl)ethan-1-one (104)**



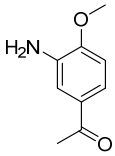
Brown solid; Yield: 61 %; mp 44-46 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 7.36 – 7.34 (m, 2H), 6.74 (d, $J$ = 8.9 Hz, 1H), 3.94 (br, 2H), 3.90 (d, $J$ = 6.9 Hz, 2H), 2.52 (s, 3H), 1.34 – 1.25 (m, 1H), 0.72 – 0.56 (m, 2H), 0.40 – 0.32 (m, 2H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 197.3, 150.7, 136.3, 130.6, 120.4, 114.0, 110.2, 73.2, 26.3, 10.2, 3.2; IR (neat): $\tilde{\upsilon}$ = 3471, 3359, 3082, 2995, 2925, 1659, 1614, 1582, 1512, 1444, 1409, 1351, 1299, 1250, 1216, 1153, 1058, 1020, 1002, 979, 941, 877, 837, 794 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{16}NO_2^+$, 206.1176; found, 206.1175.

**1-(3-Amino-4-(benzyloxy)phenyl)ethan-1-one (105)**



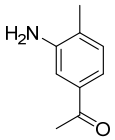
Yellow solid; Yield: 53 %; mp 119-122 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 7.44 – 7.34 (m, 7H), 6.87 (d, $J$ = 9.0 Hz, 1H), 5.16 (s, 2H), 3.93 (br, 2H), 2.52 (s, 3H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 197.2, 150.3, 136.4, 136.3, 130.9, 128.7, 128.3, 127.5, 120.3, 114.1, 110.7, 70.5, 26.3; IR (neat): ῦ = 3457, 3355, 2997, 2921, 1659, 1614, 1580, 1510, 1444, 1386, 1358, 1301, 1248, 1211, 1160, 1065, 1023, 997, 981, 927, 891, 857, 792, 749, 705 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{15}H_{16}NO_2^+$, 242.1176; found, 242.1176.

**1-(3-Amino-4-methyoxyphenyl)ethan-1-one (106)**



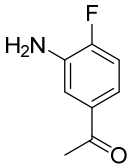
Light yellow solid; Yield: 48 %; mp 96-99 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 7.40 (dd, $J$ = 8.4, 2.2 Hz, 1H), 7.36 (d, $J$ = 2.1 Hz, 1H), 6.81 (d, $J$ = 8.3 Hz, 1H), 3.92 (s, 3H), 2.53 (s, 3H); $^{13}$C NMR (75 MHz, $CDCl_3$): δ = 197.3, 151.3, 136.1, 130.7, 120.5, 114.0, 109.2, 55.6, 26.2; IR (neat): ῦ = 3454, 3370, 2935, 1667, 1584, 1514, 1440, 1298, 1219, 905, 729, 648, 406 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_9H_{11}NNaO_2^+$, 188.1; found, 187.9.

**1-(3-Amino-4-methylphenyl)ethan-1-one (107)**

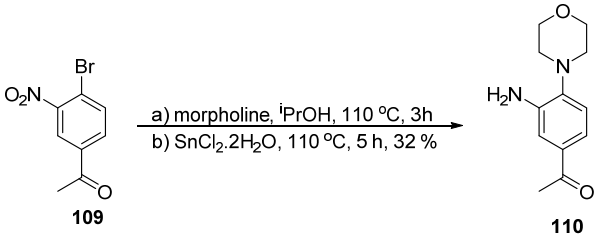


Light brown solid; Yield: 79 %; mp 77-78 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 7.30 (d, $J$ = 1.6 Hz, 1H), 7.28 (s, 1H), 7.12 (d, $J$ = 7.5 Hz, 1H), 3.77 (br, 2H), 2.54 (s, 3H), 2.22 (s, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 198.2, 144.8, 136.3 130.4, 128.0, 119.0, 113.8, 26.5, 17.5; IR (neat): ῦ = 3433, 3347, 3229, 1664, 1633, 1602, 1567, 1417, 1357, 1303, 1287, 1237, 1199, 1141, 955, 856, 834 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_9H_{12}NO^+$, 150.1; found, 149.9.

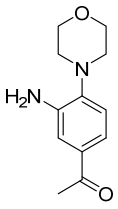
**1-(3-Amino-4-fluorophenyl)ethan-1-one (108)**



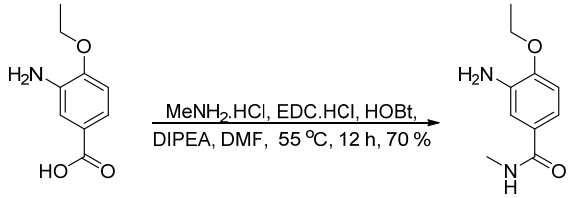
White solid; Yield: 69 %; mp 65-68 °C; $^{1}$H NMR (300 MHz, $CDCl_3$): δ= 7.41 (dd, $J$ = 8.7, 2.2 Hz, 1H) 7.31 (ddd, $J$ = 8.4, 4.6, 2.2 Hz, 1H), 7.03 (dd, $J$ = 10.6, 8.4 Hz, 1H), 3.68 (br, 2H), 2.54 (s, 3H). $^{13}$C NMR (126 MHz, $CDCl_3$): δ = 191.0, 154.5 (d, $J$ = 247.7 Hz), 134.8 (d, $J$ = 13.1 Hz), 134.0 (d, $J$ = 3.1 Hz), 119.7 (d, $J$ = 7.9 Hz), 116.5 (d, $J$ = 5.3 Hz), 115.1 (d, $J$ = 19.7 Hz), 26.4; $^{19}$F NMR (376 MHz, $CDCl_3$): δ = -127.87; IR (neat): ῦ = 3396, 3325, 3217, 1665, 1607, 1589, 1510, 1422, 1309, 1282, 1244, 1196, 1137, 1099, 1060, 964, 878, 802, 713 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_8H_8FNNaO^+$, 176.2; found, 178.8.



**1-(3-Amino-4-morpholinophenyl)ethan-1-one (110)**

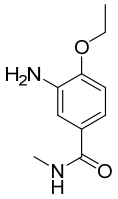


A solution of 1-(4-Bromo-3-nitrophenyl)ethanone (**109**, 500 mg, 2.05 mmol) in isopropanol (5.00 mL) was heated at 110 ºC for 3 h. $SnCl_2.2H_2O$ (1.80 g, 7.98 mmol) was then added and heated at 110 ºC for 5 h. The reaction mixture was concentrated under reduced pressure and the pH was basified to pH 5 by the addition of a 5 M NaOH solution. The resulting precipitate was filtered off and washed with EtOAc. The organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (hex: EtOAc, 1:1) affording the desired aniline in pure form as a brown solid (145 mg, 0.579 mmol, 32 % yield over two steps). mp 141-144 °C; $^1$H NMR (500 MHz, $CDCl_3$): δ = 7.37 (dd, *J* = 8.2, 1.7 Hz, 1H), 7.34 (d, *J* = 1.6 Hz, 1H), 6.99 (d, *J* = 8.1 Hz, 1H), 4.05 (br, 2H), 3.88 – 3.84 (m, 4H), 3.01 – 2.96 (m, 4H), 2.53 (s, 3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ = 197.7, 143.4, 141.0, 133.4, 120.0, 118.7, 114.5, 67.3, 50.7, 26.5; IR (neat): ῡ = 3380, 3311, 2962, 2924, 2824, 1671, 1627, 1591, 1566, 1505, 1445, 1424, 1370, 1363, 1301, 1288, 1255, 1213, 1205, 1106, 1065, 1041, 966, 941, 920, 902, 861, 845, 821, 728 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{17}N_2O_2^+$, 221.1285; found, 221.1283.




**111** **112**

**3-Amino-4-ethoxy-*N*-methylbenzamide (112)**



To a solution of 3-amino-4-ethoxybenzoic acid (**111**, 100 mg, 0.552 mmol) in DMF (1.80 mL), methylamine hydrochloride (150 mg, 2.22 mmol), HOBt (89.4 mg, 0.662 mmol), EDC.HCl (1.104 mmol, 212 mg) and DIPEA (385 µL, 2.21 mmol) were added. The solution was stirred at 55 °C for 12 h, it was concentrated and purified by flash column chromatography (5 % MeOH in DCM) affording the desired amide as a white solid in pure form (75.0 mg, 0.384 mmol, 70 % yield). mp 142-145 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 7.18 (d, *J* = 2.2 Hz, 1H), 7.06 (dd, *J* = 8.3, 2.2 Hz, 1H), 6.75 (d, *J* = 8.3 Hz, 1H), 5.98 (br, 1H), 4.10 (q, *J* = 7.0 Hz, 2H), 3.90 (br, 2H), 2.98 (d, *J* = 4.9 Hz, 3H), 1.45 (t, *J* = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 168.2, 149.0, 136.3, 127.3, 117.0, 113.6, 110.4, 63.9, 26.7, 14.8; R (neat): ῡ = 3461, 3361, 3282, 2985, 2930, 1610, 1593, 1577, 1550, 1508, 1473, 1390, 1314, 1282, 1223, 1146, 1112, 1040, 886, 823, 777 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{10}H_{15}N_2O_2^+$, 195.1128; found, 195.1129.

### III. Amide coupling reactions

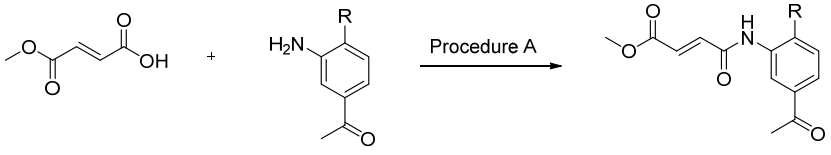
**General procedure A for amide formation**

The desired aniline (1.0 eq) and carboxylic acid (1.2 or 1.3 eq) were dissolved in DMF (0.10 M) and EDC.HCl (1.5-2.0 eq), DIPEA (1.5 eq) and HOBt (1.5-2 eq) were added at 25 ºC. In the case of **46**, **55**, **115**, **116**, **120, 132** and **133**, EDC (1.5-2.0 eq) and no DIPEA was used. The reaction mixture was stirred for 12-48 h at 25 ºC, it was concentrated and redissolved in EtOAc. The organic phase was extracted with saturated $NaHCO_3$ solution, 1M HCl and brine. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (using 1:1 hex: EtOAc, EtOAc or 2:1 EtOAc:hex as eluent), affording the desired amides in pure form. This method was used to obtain **46**, **54-56**, **115**, **116, 118**, **120**, **129** and **131-134**.

**General procedure B for amide formation**

To a solution of the carboxylic acid (1.0 eq) in toluene (1.0 M) thionyl chloride (2.0 eq) and one drop of DMF were added. The solution was refluxed for 3 h, concentrated and dissolved in DCM (0.50 M). The corresponding aniline (1.2 eq) was added and the reaction mixture was stirred at 25 °C for 12h. The reaction mixture was concentrated and purified by flash column chromatography (hex:EtOAc, 2:1) obtaining the desired amides in pure form. This method was used to obtain intermediates **123-128** and **133**.

**Synthesis of fumaric acid derivatives**

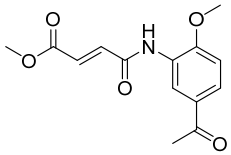



**113** (commercial)

**106**, R = methoxy
**107**, R = methyl
**108**, R = fluoro
**114**, R = H (commercially available)

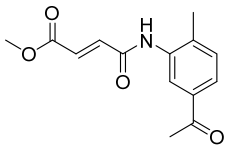
**115**, R = methoxy
**116**, R = methyl
**117**, R = fluoro
**118**, R = H

**Methyl (*E*)-4-((5-acetyl-2-methoxyphenyl)amino)-4-oxobut-2-enoate (115)**



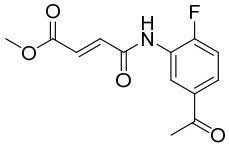
White solid; Yield: 35 %; mp 191-194 °C; $^1$H NMR (300 MHz, $CDCl_3$): δ = 9.11 (d, *J* = 1.8 Hz, 1H), 8.05 (s, 1H), 7.81 (dd, *J* = 8.6, 2.2 Hz, 1H), 7.14 – 6.91 (m, 3H), 3.99 (s, 3H), 3.84 (s, 3H), 2.60 (s, 3H); $^{13}$C NMR (126 MHz, $CDCl_3$): δ = 197.0, 165.8, 161.4, 151.5, 136.6, 131.2, 130.6, 126.8, 125.4, 120.7, 109.7, 56.2, 52.3, 26.5; IR (neat): ῦ = 3379, 2959, 1718, 1686, 1676, 1645, 1591, 1537, 1493, 1423, 1366, 1312, 1298, 1271, 1257, 1159, 1130, 1021, 976, 885, 824, 801, 763 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_{16}NO_5^+$, 278.1023; found, 278.1022.

**Methyl (*E*)-4-((5-acetyl-2-methylphenyl)amino)-4-oxobut-2-enoate (116)**



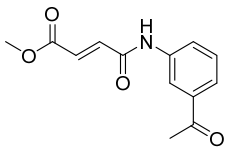
White solid; Yield: 57%; mp 130-133 °C; $^{1}$H NMR (300 MHz, $CDCl_3$): δ = 8.64 (s, 1H), 8.45 (s, 1H), 8.01 (d, *J* = 7.7 Hz, 1H), 7.63 – 7.49 (m, 2H), 7.35 – 7.15 (m, 1H), 4.14 (s, 3H), 2.88 (s, 3H), 2.88 (s, 3H); $^{13}$C NMR (100 MHz, DMSO-*d*$_6$): δ = 197.1, 165.4, 161.7, 137.6, 137.0, 135.9, 135.1, 130.8, 129.3, 125.4, 123.8, 52.1, 26.6, 18.0; IR (neat): ῡ = 3353, 2954, 1715, 1685, 1670, 1652, 1605, 1529, 1499, 1439, 1355, 1316, 1274, 1226, 1201, 1161, 1139, 1016, 1003, 985, 822, 795, 768 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_{16}NO_4^+$, 262.1074; found, 262.1072.

**Methyl (*E*)-4-((5-acetyl-2-fluorophenyl)amino)-4-oxobut-2-enoate (117)**

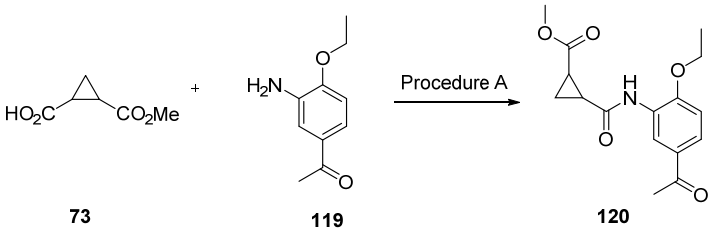


To a solution of mono methyl fumarate (**113**, 127 mg, 0.980 mmol) at 0 °C in DCM (1.00 mL) oxalyl chloride (84.1 µL, 0.980 mmol) and one drop of DMF were added. The solution was stirred at 0 ºC for 20 min and at 25 °C for 30 min. The aniline, 1-(3-amino-4-fluorophenyl)ethan-1-one (**72**, 50.0 mg, 0.327 mmol), was slowly added at 0 ºC, followed by $Et_3N$ (209 µL, 1.50 mmol) and DMAP (12.0 mg, 0.0980 mmol). The reaction mixture was stirred at 25 ºC for 12 h, it was quenched with a saturated solution of $NH_4Cl$ and extracted with EtOAc three times. The combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (1:1 hex: EtOAc) affording the desired amide in pure form as a white solid (37.0 mg, 0.140 mmol, 43 % yield). mp 176-179 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 9.06 (d, *J* = 6.7 Hz, 1H), 7.79 (ddd, *J* = 8.5, 5.1, 2.2 Hz, 2H), 7.21 (dd, *J* = 10.3, 8.7 Hz, 1H), 7.13 (d, *J* = 15.3 Hz, 1H), 7.00 (d, *J* = 15.3 Hz, 1H), 3.85 (s, 3H), 2.62 (s, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 196.5, 165.6, 161.6, 135.7, 134.1 (d, *J* = 3.2 Hz), 132.1, 126.0 (d, *J* = 10.8 Hz), 125.6 (d, *J* = 8.2 Hz), 122.6, 115.4, 115.2, 52.42, 26.63; IR (neat): ῡ = 3354, 2963, 2924, 1715, 1686, 1650, 1614, 1600, 1540, 1487, 1441, 1418, 1357, 1320, 1293, 1261, 1215, 1183, 1153, 1113, 974, 877, 818, 766 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{13}FNO_4^+$, 266.0823; found, 266.0823.

**Methyl (*E*)-4-((3-acetylphenyl)amino)-4-oxobut-2-enoate (118)**

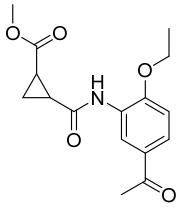


White solid; Yield: 50 %; mp 177-178 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 8.64 (s, 1H), 8.21 (d, *J* = 7.5 Hz, 1H), 8.14 (s, 1H), 7.75 (d, *J* = 7.7 Hz, 1H), 7.49 (t, *J* = 7.9 Hz, 1H), 7.28 (d, *J* = 14.3 Hz, 1H), 7.01 (d, *J* = 15.3 Hz, 1H), 3.87 (s, 3H), 2.67 (s, 3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ = 198.4, 166.2, 161.8, 138.4, 137.6, 136.9, 131.2, 129.6, 125.0, 124.7, 119.2, 52.5, 26.7; IR (neat): ῡ = 3332, 1723, 1687, 1673, 1607, 1551, 1485, 1440, 1360, 1334, 1299, 1275, 1223, 1204, 1154, 1003, 973, 898, 878, 811, 711 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{14}NO_4^+$, 248.0917; found, 248.09144.



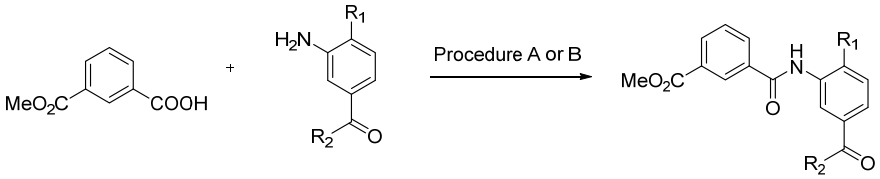

**73** **119** **120**

**Methyl 2-((5-acetyl-2-ethoxyphenyl)carbamoyl)cyclopropane-1-carboxylate (120)**



White solid; Yield: 55 %; mp 166-168 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 8.99 (s, 1H), 8.05 (s, 1H), 7.74 (dd, $J$ = 8.6, 2.1 Hz, 1H), 6.92 (d, $J$ = 8.7 Hz, 1H), 4.21 (q, $J$ = 7.0 Hz, 2H), 3.74 (s, 3H), 2.56 (s, 3H), 2.35 – 2.26 (m, 1H), 2.20 – 2.11 (m, 1H), 1.63 – 1.57 (m, 1H), 1.52 (t, $J$ = 7.0 Hz, 3H), 1.49 – 1.44 (m, 1H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.0, 173.0, 168.4, 150.5, 130.4, 127.2, 124.6, 120.4, 110.4, 64.7, 52.2, 26.5, 25.5, 22.0, 15.5, 14.7; IR (neat): ῦ = 3289, 2986, 2951, 2925, 1723, 1682, 1653, 1601, 1586, 1544, 1500, 1423, 1362, 1344, 1277, 1231, 1203, 1172, 1131, 1039, 956, 925, 875, 815, cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{16}H_{19}NNaO_5^+$, 328.1155; found, 328.1152.

**Synthesis of benzoic acid derivatives. Scope on the acyl benzene**

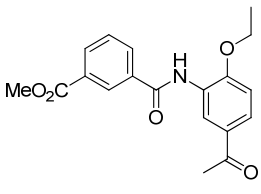



**121**

**119**, $R_1$ = OEt, $R_2$ = Me (commercially available)
**122**, $R_1$ = Br, $R_2$ = Me (commercially available)
**102**, $R_1$ = OPr, $R_2$ = Me
**103**, $R_1$ = O$^{i}$Bu, $R_2$ = Me
**104**, $R_1$ =  , $R_2$ = Me
**105**, $R_1$ = OBenzyl, $R_2$ = Me
**110**, $R_1$ = morpholine, $R_2$ = Me
**112**, $R_1$ = OEt, $R_2$ = NHMe

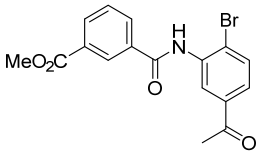
**46**, $R_1$ = OEt, $R_2$ = Me
**123**, $R_1$ = Br, $R_2$ = Me
**124**, $R_1$ = OPr, $R_2$ = Me
**125**, $R_1$ = O$^{i}$Bu, $R_2$ = Me
**126**, $R_1$ =  , $R_2$ = Me
**127**, $R_1$ = OBenzyl, $R_2$ = Me
**128**, $R_1$ = morpholine, $R_2$ = Me
**129**, $R_1$ = OEt, $R_2$ = NHMe

**Methyl 3-((5-acetyl-2-ethoxyphenyl)carbamoyl)benzoate (46)**



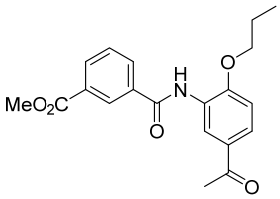
White solid; Yield: 80%; mp 137-140 °C; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 9.14 (d, $J$ = 2.2 Hz, 1H), 8.65 (s, 1H), 8.53 (t, $J$ = 1.5 Hz, 1H), 8.28 – 8.18 (m, 1H), 8.12 (ddd, $J$ = 7.8, 1.9, 1.2 Hz, 1H), 7.77 (dd, $J$ = 8.6, 2.2 Hz, 1H), 7.60 (td, $J$ = 7.8, 0.5 Hz, 1H), 6.96 (d, $J$ = 8.6 Hz, 1H) 4.24 (q, $J$ = 7.0 Hz, 2H), 3.96 (s, 3H), 2.60 (s, 3H), 1.54 (t, $J$ = 7.0 Hz, 3H);$^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.0, 166.1, 164.1, 151.1, 135.1, 132.8, 131.5, 130.8, 130.4, 129.2, 127.7, 127.2, 124.9, 120.5, 110.4, 64.8, 52.4, 26.5, 14.7; IR (neat): ῦ = 3430, 2979, 2954, 1720, 1678, 1590, 1537, 1468, 1433, 1294, 1261, 1228, 1205, 1145, 1033, 988, 907, 797, 728 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{20}NO_5^+$, 342.1336; found, 342.1331.

**Methyl 3-((5-acetyl-2-bromophenyl)carbamoyl)benzoate (123)**



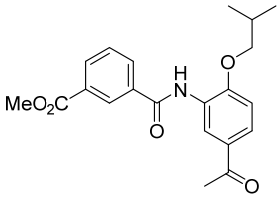
White solid; Yield: 60 %; mp 133-135 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 9.14 (d, $J$ = 1.9 Hz, 1H), 8.61 (s, 1H), 8.54 (br, 1H), 8.28 (dd, $J$ = 7.8, 1.3 Hz, 1H), 8.17 (d, $J$ = 7.8 Hz, 1H), 7.75 – 7.60 (m, 3H), 3.98 (s, 3H), 3.98 (s, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.1, 166.0, 164.5, 137.4, 135.9, 134.5, 133.3, 132.7, 131.6, 131.1, 129.4, 128.0, 124.7, 123.0, 119.0, 52.5, 26.7; IR (neat): ῡ = 3242, 2923, 2853, 1726, 1691, 1649, 1572, 1524, 1436, 1414, 1315, 1295, 1245, 1213, 1092, 1083, 1027, 932, 817, 732, 724, 706 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{17}H_{14}BrNNaO_4^+$, 397.9998; found, 397.9996.

**Methyl 3-((5-acetyl-2-propoxyphenyl)carbamoyl)benzoate (124)**



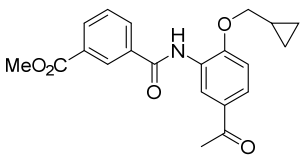
White solid; Yield: 40 %; mp 132-135 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 9.18 (d, $J$ = 2.1 Hz, 1H), 8.69 (s, 1H), 8.54 (t, $J$ = 1.5 Hz, 1H), 8.28 – 8.21 (m, 1H), 8.15 (ddd, $J$ = 7.8, 1.9, 1.2 Hz, 1H), 7.80 (dd, $J$ = 8.6, 2.2 Hz, 1H), 7.62 (t, $J$ = 7.8 Hz, 1H), 6.98 (d, $J$ = 8.6 Hz, 1H), 4.15 (t, $J$ = 6.5 Hz, 2H), 3.97 (s, 3H), 2.63 (s, 3H), 2.03 – 1.87 (m, 2H), 1.14 (t, $J$ = 7.4 Hz, 3H); $^{13}$C NMR (126 MHz, $CDCl_3$): 197.2, 166.1, 164.1, 151.2, 135.1, 132.9, 131.7, 130.9, 130.5, 129.3, 127.6, 127.3, 124.9, 120.5, 110.5, 70.6, 52.4, 26.6, 22.5, 10.5; IR (neat): ῡ = 3321, 2970, 2923, 1726, 1681, 1655, 1602, 1585, 1537, 1498, 1461, 1422, 1361, 1331, 1300, 1272, 1202, 1137, 1096, 1074, 975, 887, 805, 728, 719 cm$^{-1}$;HRMS (ESI), *m/z*: calcd for $C_{20}H_{22}NO_5^+$, 356.1493; found, 356.1493.

**Methyl 3-((5-acetyl-2-isobutoxyphenyl)carbamoyl)benzoate (125)**



Beige solid; Yield: 89 %; mp 116-119 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 9.19 (d, $J$ = 2.1 Hz, 1H), 8.71 (s, 1H), 8.54 (d, $J$ = 1.6 Hz, 1H), 8.29 – 8.21 (m, 1H), 8.19 – 8.12 (m, 1H), 7.79 (dd, $J$ = 8.6, 2.2 Hz, 1H), 7.62 (t, $J$ = 7.8 Hz, 1H), 6.98 (d, $J$ = 8.6 Hz, 1H), 3.94 – 3.96 (m, 5H), 2.62 (s, 3H), 2.24 (dp, $J$ = 13.3, 6.7 Hz, 1H), 1.13 (d, $J$ = 6.7 Hz, 6H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.1, 166.0, 164.0, 151.2, 135.1, 132.9, 131.7, 130.9, 130.5, 129.3, 127.4, 127.3, 124.9, 120.3, 110.5, 75.2, 52.4, 28.3, 26.6, 19.2; IR (neat): ῡ = 3317, 2959, 1725, 1678, 1652, 1601, 1583, 1535, 1497, 1468, 1422, 1360, 1331, 1301, 1277, 1258, 1201, 1140, 1094, 1075, 1018, 885, 809, 727 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{21}H_{24}NO_5^+$, 370.1649; found, 370.1649.

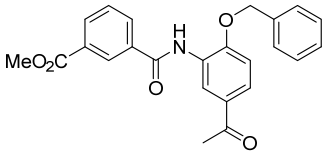
**Methyl 3-((5-acetyl-2-(cyclopropylmethoxy)phenyl)carbamoyl)benzoate (126)**



Pale yellow solid; Yield: 91 %; mp 141-142 °C; $^{1}$H NMR (400 MHz, $CDCl_3$): δ = 9.17 (d, $J$ = 2.1 Hz, 1H), 8.78 (s, 1H), 8.56 (t, $J$ = 1.5 Hz, 1H), 8.27 – 8.21 (m, 1H), 8.21 – 8.13 (m, 1H), 7.77 (dd, $J$ = 8.6,

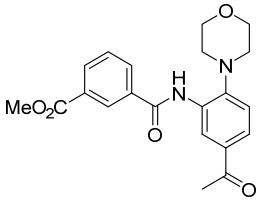
2.2 Hz, 1H), 7.62 (t, $J$ = 7.8 Hz, 1H), 6.94 (d, $J$ = 8.6 Hz, 1H), 4.02 (d, $J$ = 7.1 Hz, 2H), 3.96 (s, 3H), 2.61 (s, 3H), 1.44 – 1.30 (m, 1H), 0.78 – 0.66 (m, 2H), 0.46 – 0.42 (m, 2H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.0, 166.0, 164.0, 151.2, 135.1, 132.8, 131.7, 130.8, 130.5, 129.2, 127.6, 127.5, 124.8, 120.4, 110.8, 74.0, 52.4, 26.5, 10.1, 3.2; IR (neat): ῦ = 3426, 2923, 2853, 1722, 1672, 1587, 1531, 1473, 1462, 1432, 1340, 1305, 1295, 1267, 1254, 1228, 1204, 1146, 1098, 1077, 1026, 990, 921, 905, 804, 729, 720 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{21}H_{22}NO_5^+$, 368.1493; found, 368.1495.

**Methyl 3-((5-acetyl-2-(benzyloxy)phenyl)carbamoyl)benzoate (127)**



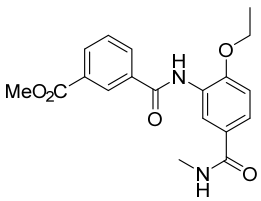
White solid; Yield: 74 %; mp 149-152 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 9.20 (d, $J$ = 2.1 Hz, 1H), 8.67 (s, 1H), 8.51 (d, $J$ = 1.6 Hz, 1H), 8.26 – 8.18 (m, 1H), 8.07 (ddd, $J$ = 7.7, 1.8, 1.2 Hz, 1H), 7.80 (dd, $J$ = 8.6, 2.2 Hz, 1H), 7.57 (t, $J$ = 7.7 Hz, 1H), 7.48 – 7.37 (m, 5H), 7.08 (d, $J$ = 8.6 Hz, 1H), 5.27 (s, 2H), 3.94 (s, 3H), 2.63 (s, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.0, 166.0, 164.1, 150.9, 135.5, 135.0, 132.8, 131.5, 130.9, 130.8, 129.1, 128.9, 128.6, 127.7, 127.6, 127.3, 124.8, 120.6, 111.3, 71.2, 52.4, 26.5; IR (neat): ῦ = 3303, 3064, 2954, 2921, 1732, 1717, 1681, 1651, 1602, 1585, 1536, 1498, 1425, 1332, 1282, 1202, 1146, 1078, 1024, 973, 885, 832, 797, 728, 719 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{24}H_{21}NNaO_5^+$, 426.1312; found, 426.1307.

**Methyl 3-((5-acetyl-2-morpholinophenyl)carbamoyl)benzoate (128)**



Pale yellow solid; Yield: 95 %; mp 152-155 °C; $^1$H NMR (500 MHz, $CDCl_3$): δ = 9.51 (s, 1H), 9.17 (s, 1H), 8.58 (d, $J$ = 1.1 Hz, 1H), 8.26 (t, $J$ = 8.7 Hz, 2H), 7.81 (d, $J$ = 8.3 Hz, 1H), 7.66 (t, $J$ = 7.7 Hz, 1H), 7.31 (d, $J$ = 8.3 Hz, 1H), 4.00 (s, 3H), 3.99-3.97 (m, 4H), 3.01 (s, 4H), 2.65 (s, 3H); $^{13}$C NMR (126 MHz, $CDCl_3$): δ = 197.5, 166.0, 163.7, 145.3, 134.8, 134.6, 133.2, 133.0, 132.0, 130.8, 129.5, 127.1, 124.3, 120.6, 120.1, 67.5, 52.5, 52.3, 26.7; IR (neat): ῦ = 3327, 2958, 2582, 1731, 1673, 1579, 1525, 1427, 1356, 1292, 1258, 1234, 1217, 1109, 932, 918, 814, 801, 719 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{21}H_{23}N_2O_5^+$, 383.1602; found, 383.1602.

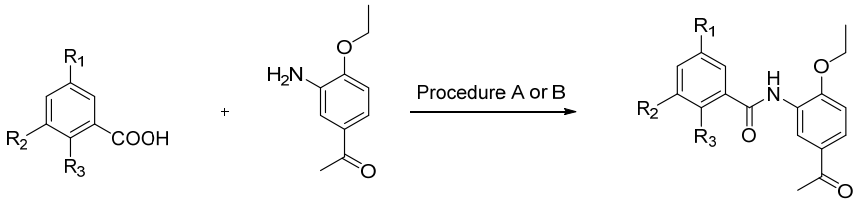
**Methyl 3-((2-ethoxy-5-(methylcarbamoyl)phenyl)carbamoyl)benzoate (129)**



White solid; Yield: 66 %; mp 161-163 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 8.81 (d, $J$ = 2.1 Hz, 1H), 8.65 (s, 1H), 8.49 (s, 1H), 8.20 (d, $J$ = 7.8 Hz, 1H), 8.12 – 8.01 (m, 1H), 7.70 (dd, $J$ = 8.5, 2.1 Hz, 1H), 7.57 (t, $J$ = 7.8 Hz, 1H), 7.57 (t, $J$ = 7.8 Hz, 1H), 6.45 (d, $J$ = 4.2 Hz, 1H), 4.17 (q, $J$ = 7.0 Hz, 2H), 3.94 (s, 3H), 2.97 (d, $J$ = 4.8 Hz, 3H), 1.50 (t, $J$ = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 167.5, 166.0, 164.1, 149.7, 135.0, 132.7, 131.4, 130.8, 129.1, 127.7, 127.1, 126.9, 124.7, 117.1, 110.7, 64.6, 52.4, 26.7, 14.6; IR (neat): ῦ = 3432, 3305,

2935, 1723, 1683, 1655, 1590, 1537, 1490, 1467, 1437, 1316, 1302, 1260, 1235, 1216, 1143, 1113, 1040, 1023, 974, 902, 815, 802, 718 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{20}N_2NaO_5^+$, 379.1264; found, 379.1262.

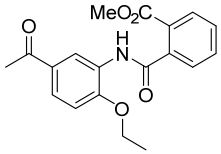
**Synthesis of benzoic acid derivatives. Scope on the side chain**




**130**, $R_1$ = H, $R_2$ = H, $R_3$ = $CO_2Me$ (commercially available) **119**
**75**, $R_1$ = H, $R_2$ = tetrazole, $R_3$ = H
**80**, $R_1$ = H, $R_2$ = N-phenylsulfamide, $R_3$ = H
**82**, $R_1$ = H, $R_2$ = phenylsulfonamide, $R_3$ = H
**84**, $R_1$ = Br, $R_2$ = $CO_2Me$, $R_3$ = H
**89**, $R_1$ = $CH_2OMe$, $R_2$ = $CO_2Me$, $R_3$ = H
**90**, $R_1$ = $CH_2OTHP$, $R_2$ = $CO_2Me$, $R_3$ = H

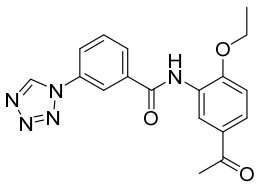
**131**, $R_1$ = H, $R_2$ = H, $R_3$ = $CO_2Me$
**54**, $R_1$ = H, $R_2$ = tetrazole, $R_3$ = H
**55**, $R_1$ = H, $R_2$ = N-phenylsulfamide, $R_3$ = H
**56**, $R_1$ = H, $R_2$ = phenylsulfonamide, $R_3$ = H
**132**, $R_1$ = Br, $R_2$ = $CO_2Me$, $R_3$ = H
**133**, $R_1$ = $CH_2OMe$, $R_2$ = $CO_2Me$, $R_3$ = H
**134**, $R_1$ = $CH_2OTHP$, $R_2$ = $CO_2Me$, $R_3$ = H

**Methyl 2-((5-acetyl-2-ethoxyphenyl)carbamoyl)benzoate (131)**



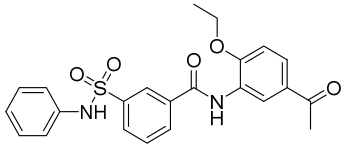

Pale brown solid; Yield: 40 %; mp 130-134 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 9.16 (s, 1H), 8.09 (s, 1H), 7.96 (d, *J* = 7.9 Hz, 1H), 7.78 (dd, *J* = 8.6, 1.8 Hz, 1H), 7.63 – 7.54 (m, 3H), 6.94 (d, *J* = 8.6 Hz, 1H), 4.19 (q, *J* = 7.0 Hz, 2H), 3.87 (s, 3H), 2.62 (s, 3H), 1.44 (t, *J* = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.2, 167.1, 167.0, 150.9, 138.2, 132.2, 130.5, 130.3, 130.1, 129.3, 127.5, 127.5, 124.8, 120.7, 110.5, 64.7, 52.7, 26.6, 14.6; IR (neat): ῡ = 3274, 2954, 2926, 1727, 1670, 1651, 1600, 1584, 1537, 1497, 1473, 1427, 1357, 1335, 1265, 1202, 1141, 1115, 1071, 1035, 959, 805, 767, 709 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{19}NNaO_5^+$, 364.1155; found, 364.1155.

***N*-(5-Acetyl-2-ethoxyphenyl)-3-(1*H*-tetrazol-1-yl)benzamide (54)**




Brown solid; Yield: 98 %; mp 168-169 °C; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.19 (s, 1H), 9.85 (s, 1H), 8.49 – 8.48 (m, 1H), 8.31 (d, *J* = 2.2 Hz, 1H), 8.17 – 8.11 (m, 2H), 7.87 (dd, *J* = 8.6, 2.2 Hz, 1H), 7.84 (t, *J* = 8.0 Hz, 1H), 7.22 (d, *J* = 8.7 Hz, 1H), 4.22 (q, *J* = 7.0 Hz, 2H), 2.55 (s, 3H), 1.38 (t, *J* = 6.9 Hz, 3H); $^{13}$C NMR (126 MHz, $CDCl_3$): δ = 196.2, 164.0, 155.2, 142.5, 136.2, 133.9, 130.5, 129.3, 128.5, 127.6, 126.5, 124.9, 124.1, 120.4, 111.9, 64.4, 26.4, 14.4; IR (neat): ῡ = 3439, 3086, 2970, 2923, 1666, 1591, 1535, 1489, 1430, 1363, 1336, 1260, 1242, 1203, 1133, 1090, 1078, 1032, 1006, 963, 893, 808, 742, 732 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{18}H_{18}N_5O_3^+$, 352.1404; found, 352.1404.

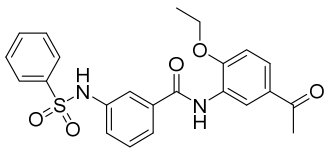
***N*-(5-Acetyl-2-ethoxyphenyl)-3-(*N*-phenylsulfamoyl)benzamide (55)**




White solid; Yield: 63 %; mp 186-189 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 10.41 (s, 1H), 9.85 (s, 1H), 8.34 (s, 1H), 8.26 (d, *J*

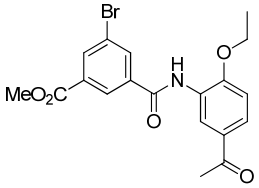
= 1.9 Hz, 1H), 8.17 (d, *J* = 7.7 Hz, 1H), 7.93 (d, *J* = 8.1 Hz, 1H), 7.86 (dd, *J* = 8.6, 2.1 Hz, 1H), 7.71 (t, *J* = 7.8 Hz, 1H), 7.31 – 7.16 (m, 3H), 7.12 (s, 1H), 7.11 (s, 1H), 7.03 (t, *J* = 7.4 Hz, 1H), 4.19 (q, *J* = 6.9 Hz, 2H), 2.53 (s, 3H), 1.35 (t, *J* = 6.9 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-*d*$_6$): δ = 196.2, 163.9, 155.3, 140.1, 137.4, 135.3, 131.5, 129.7, 129.5, 129.3, 129.2, 127.6, 126.5, 126.2, 124.9, 124.3, 120.2, 111.9, 64.4, 26.4, 14.4; IR (neat): ῡ = 3423, 3235, 2986, 1682, 1667, 1600, 1589, 1536, 1482, 1469, 1433, 1345, 1258, 1228, 1211, 1159, 1111, 1029, 915, 890, 814, 764, 754, 740 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{23}H_{23}N_2O_5S^+$, 439.1322; found, 439.1316.

**N-(5-Acetyl-2-ethoxyphenyl)-3-(phenylsulfonamido)benzamide (56)**



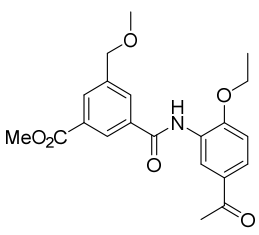
Pale yellow solid; Yield: 52 % over two steps; mp 217-220 °C; $^1$H NMR (500 MHz, $CDCl_3$): δ = 10.56 (br, 1H), 9.44 (s, 1H), 8.37 (d, *J* = 2.0 Hz, 1H), 7.83 (dd, *J* = 8.7, 2.1 Hz, 1H), 7.79 (d, *J* = 7.4 Hz, 2H), 7.68 (s, 1H), 7.63 – 7.60 (m, 2H), 7.58-7.55 (m, 2H), 7.40 (t, *J* = 7.8 Hz, 1H), 7.34 – 7.28 (m, 1H), 7.19 (d, *J* = 8.7 Hz, 1H), 4.19 (q, *J* = 7.0 Hz, 2H), 2.53 (s, 3H), 1.37 (t, *J* = 6.9 Hz, 3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ = 196.2, 164.5, 154.5, 139.3, 138.1, 135.4, 133.1, 129.5, 129.4, 127.1, 126.8, 126.6, 123.6, 122.9, 122.7, 119.1, 111.7, 64.4, 26.4, 14.4, 1 C is missing due to overlapping; IR (neat): ῡ = 3393, 3168, 2973, 2908, 1667, 1598, 1584, 1530, 1472, 1423, 1366, 1332, 1273, 1156, 1139, 1089, 1044, 971, 899, 881, 808, 743, 720 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{23}H_{23}N_2O_5S^+$, 439.1322; found, 439.1324.

**Methyl 3-((5-acetyl-2-ethoxyphenyl)carbamoyl)-5-bromobenzoate (132)**



Pale yellow solid; Yield: 44 % over two steps; mp 159-161 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 9.09 (d, *J* = 2.2 Hz, 1H), 8.58 (br, 1H), 8.43 (t, *J* = 1.5 Hz, 1H), 8.38 – 8.33 (m, 1H), 8.27 (t, *J* = 1.8 Hz, 1H), 7.80 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.98 (d, *J* = 8.7 Hz, 1H), 4.26 (q, *J* = 7.0 Hz, 2H), 3.98 (s, 3H), 2.62 (s, 3H), 1.55 (t, *J* = 7.0 Hz, 3H);$^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.0, 164.8, 162.7, 151.2, 136.9, 135.7, 134.8, 132.5, 130.5, 126.9, 126.1, 125.3, 123.3, 120.6, 110.5, 64.9, 52.7, 26.5, 14.7; IR (neat): ῡ = 3414, 2960, 2922, 1739, 1689, 1673, 1588, 1531, 1485, 1437, 1340, 1290, 1259, 1227, 1205, 1150, 1111, 1070, 1029, 981, 931, 889, 800, 750, 733, 723 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{18}BrNNaO_5^+$, 442.0261; found, 442.0264.

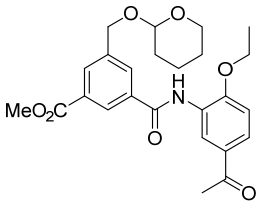
**Methyl 3-((5-acetyl-2-ethoxyphenyl)carbamoyl)-5-(methoxymethyl)benzoate (133)**



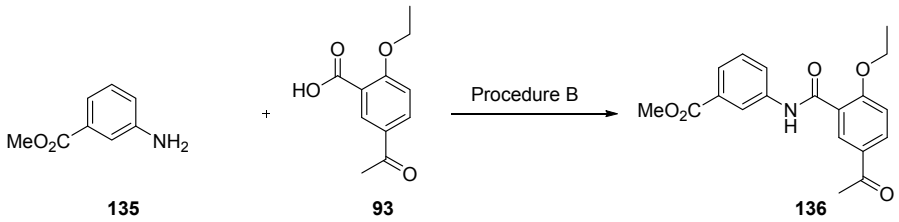
White solid; Yield: 17 % over two steps; mp 123-125 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 9.16 (d, *J* = 2.1 Hz, 1H), 8.67 (s, 1H), 8.47 (s, 1H), 8.21 (s, 1H), 8.13 (s, 1H), 8.13 (s, 1H), 6.98 (d, *J* = 8.7 Hz, 1H), 4.59 (s, 2H), 4.26 (q, *J* = 7.0 Hz, 2H), 3.97 (s, 3H), 3.46 (s, 3H), 2.63 (s, 3H), 1.56 (t, *J* = 6.9 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.1, 166.1, 164.1, 151.2, 140.2, 135.4, 131.7, 131.1, 130.5, 127.2, 126.9, 124.9, 120.6, 110.5, 73.6, 64.8, 58.6, 52.5, 26.6, 14.7, 1 C is missing due to overlapping; IR (neat): ῡ = 3422, 2922, 2851, 1724, 1686,

1669, 1601, 1588, 1531, 1485, 1431, 1262, 1223, 1206, 1146, 1112, 1031, 999, 885, 794, 746, 727 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{21}H_{24}NO_6^+$, 386.1598; found, 386.1596.

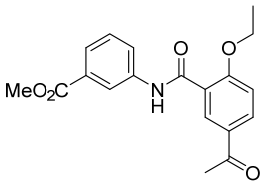
**Methyl 3-((5-acetyl-2-ethoxyphenyl)carbamoyl)-5-(((tetrahydro-2*H*-pyran-2-yl)oxy)methyl)benzoate (134)**



White solid; Yield: 44 %; mp 133-135 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 9.16 (d, *J* = 2.1 Hz, 1H), 8.65 (s, 1H), 8.46 (s, 1H), 8.22 (s, 1H), 8.15 (s, 1H), 7.79 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.98 (d, *J* = 8.6 Hz, 1H), 4.91 (d, *J* = 12.6 Hz, 1H), 4.76 (t, *J* = 3.5 Hz, 1H), 4.62 (d, *J* = 12.6 Hz, 1H), 4.25 (q, *J* = 7.0 Hz, 2H), 3.97 (s, 3H), 3.95 – 3.87 (m, 1H), 3.62 – 3.53 (m, 1H), 2.62 (s, 3H), 1.92 – 1.84 (m, 1H), 1.83 – 1.57 (m, 5H), 1.55 (t, *J* = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.1, 166.1, 164.2, 151.1, 140.2, 135.3, 131.8, 131.0, 130.6, 130.5, 127.2, 126.8, 124.9, 120.6, 110.5, 98.3, 68.0, 64.8, 62.2, 52.4, 30.5, 26.5, 25.4, 19.3, 14.7; IR (neat): ῡ = 3437, 2935, 2864, 1721, 1685, 1669, 1593, 1552, 1489, 1441, 1343, 1313, 1256, 1214, 1200, 1182, 1136, 1115, 1022, 974, 906, 871, 802, 740, 716 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{25}H_{29}NNaO_7^+$, 478.1836; found, 478.1835.

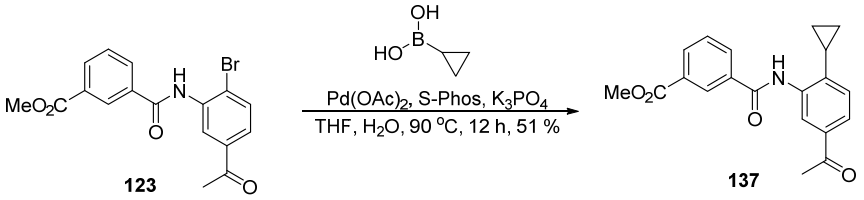



**Methyl 3-(5-acetyl-2-ethoxybenzamido)benzoate (136)**

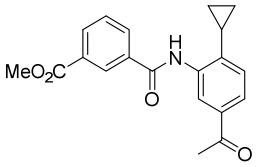


Pale yellow solid; Yield: 34 %; mp 128-130 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 10.04 (s, 1H), 8.87 (d, *J* = 2.4 Hz, 1H), 8.22 (t, *J* = 1.8 Hz, 1H), 8.14 (dd, *J* = 8.7, 2.4 Hz, 1H), 7.98 (ddd, *J* = 8.1, 2.2, 1.0 Hz, 1H), 7.81 (ddd, *J* = 7.7, 1.6, 1.1 Hz, 1H), 7.45 (t, *J* = 7.9 Hz, 1H), 7.09 (d, *J* = 8.8 Hz, 1H), 4.38 (q, *J* = 7.0 Hz, 2H), 3.93 (s, 3H), 2.63 (s, 3H), 1.69 (t, *J* = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 196.4, 166.7, 162.4, 160.0, 138.5, 133.8, 133.3, 131.0, 130.9, 129.2, 125.4, 124.4, 120.9, 120.9, 112.6, 65.7, 52.2, 26.5, 14.8; IR (neat): ῡ = 3319, 2951, 2925, 1722, 1671, 1599, 1560, 1489, 1435, 1363, 1335, 1291, 1268, 1239, 1213, 1162, 1109, 1071, 1025, 969, 884, 814, 795, 751 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{20}NO_5^+$, 342.1336; found, 342.1336.

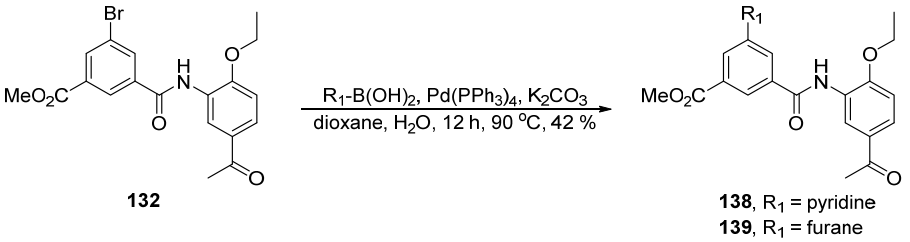
## IV. Suzuki cross-coupling reactions



**Methyl 3-((5-acetyl-2-cyclopropylphenyl)carbamoyl)benzoate (137)**



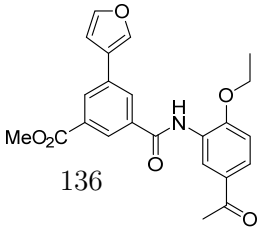
To a solution of methyl 3-((5-acetyl-2-bromophenyl)carbamoyl)benzoate (**123**, 33 mg, 0.088 mmol) in THF:$H_2O$ (3:1, 0.22 M) $Pd(OAc)_2$ (2.0 mg, 0.0089 mmol), S-Phos (4.0 mg, 0.0097 mmol) and $K_3PO_4$ (68 mg, 0.312 mmol) were added. The solution was purged with $N_2$ and heated to 90 ºC for 12 hours. The reaction mixture was diluted with water and extracted with EtOAc three times. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash column chromatography (1:1 hex:EtOAc) affording the final product in pure form as a pale brown solid (15 mg, 0.044 mmol, 51 % yield). mp 125-130 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 8.92 (d, *J* = 1.7 Hz, 1H), 8.66 (br, 1H), 8.59 (t, *J* = 1.6 Hz, 1H), 8.28 – 8.23 (m, 1H), 8.22 – 8.16 (m, 1H), 7.72 (dd, *J* = 8.1, 1.8 Hz, 1H), 7.63 (t, *J* = 7.8 Hz, 1H), 7.29 (d, *J* = 8.1 Hz, 1H), 3.97 (s, 3H), 2.63 (s, 3H), 1.93 (tdd, *J* = 10.9, 7.3, 3.6 Hz, 1H), 1.20 – 1.09 (m, 2H), 0.86 – 0.76 (m, 2H); $^{13}$C NMR (101 MHz, $CDCl_3$): δ = 197.7, 166.1, 164.2, 137.7, 136.8, 136.3, 135.0, 132.9, 131.8, 130.9, 129.4, 128.6, 127.5, 124.1, 121.0, 52.5, 26.7, 11.7, 6.2; IR (neat): $\tilde{\upsilon}$ = 3267, 2954, 2924, 1725, 1677, 1643, 1606, 1569, 1525, 1413, 1357, 1319, 1282, 1251, 1221, 1201, 1136, 1098, 1045, 1015, 987, 898, 880, 813, 797, 727 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{20}H_{19}NNaO_4^+$, 360.1206; found, 360.1203.




**General procedure for Suzuki cross-couplings**

To a solution of methyl 3-((5-acetyl-2-ethoxyphenyl)carbamoyl)-5-bromobenzoate (**132**) in dioxane (0.24 M) and water (1 drop), the corresponding boronic acid (1.1 eq), $K_2CO_3$ (3.0 eq) and $Pd(PPh_3)_4$ (0.25 eq) were added. The reaction was stirred for 12 h at 90 °C. It was diluted with water and extracted with EtOAc three times. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure. The obtained residue was purified by flash column chromatography (hex: EtOAc, 2:1 to pure EtOAc) affording the coupled products. This procedure was used to obtain intermediates **138** and **139**. **138** was extracted and used in the next step without further purification.

**Methyl 3-((5-acetyl-2-ethoxyphenyl)carbamoyl)-5-(furan-3-yl)benzoate (139)**



136

Yellow solid; Yield: 42 %; mp 156-163 °C; $^1$H NMR (400 MHz, $CDCl_3$): δ = 9.16 (d, *J* = 2.2 Hz, 1H), 8.70 (s, 1H), 8.37 (t, *J* = 1.6 Hz, 1H), 8.32 (t, *J* = 1.6 Hz, 1H), 8.27 (t, *J* = 1.8 Hz, 1H), 7.88 (dd, *J* = 1.4, 0.9 Hz,

1H), 7.80 (dd, $J$ = 8.6, 2.2 Hz, 1H), 7.54 (t, $J$ = 1.6 Hz, 1H), 6.98 (d, $J$ = 8.6 Hz, 1H), 6.81 (dd, $J$ = 1.9, 0.9 Hz, 1H), 4.26 (q, $J$ = 7.0 Hz, 2H), 3.98 (s, 3H), 2.62 (s, 3H), 1.57 (t, $J$ = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ = 197.0, 166.0, 164.0, 151.2, 144.3, 139.5, 135.7, 134.1, 131.4, 130.5, 129.9, 129.1, 127.2, 125.5, 125.0, 124.9, 120.6, 110.5, 108.6, 64.8, 52.5, 26.5, 14.7; IR (neat): ῡ = 2927, 2851, 1724, 1682, 1592, 1535, 1434, 1261, 1024, 803, 794, 765, 748 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{23}H_{21}NNaO_6^+$, 430.1261; found, 430.1260.

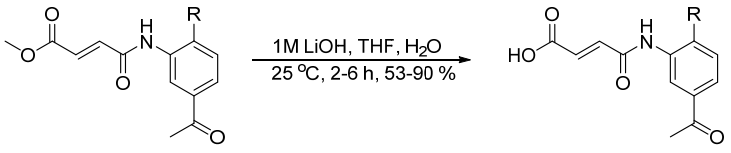
### V. Methyl ester hydrolysis of the amide coupling products

#### General procedure for ester hydrolysis

To a solution of the methyl ester (1 eq) in THF (0.1 M) a 1M LiOH solution (5 eq) was added. The reaction mixture was stirred at 25 °C for 2-6 h. It was then concentrated under reduced pressure and 1M HCl was added. The obtained precipitate was washed with hexanes, $Et_2O$ and cold DCM, affording the desired carboxylic acids in pure form. For the synthesis of **57**, the reaction mixture was diluted with water and extracted with DCM and $Et_2O$ three times. The pH of the water phase was then brought to pH 3-4 by addition of 1M HCl and the resulting precipitate was filtered off and washed with DCM.

For the synthesis of **59**, the pH of the reaction mixture was brought to 1 by the addition of 1M HCl solution. It was then concentrated, redissolved in EtOH (0.1 M) and PTSA (0.1 eq) was added. The solution was stirred at 25 °C for 5 h and concentrated. Upon the addition of 1M HCl solution, the final product precipitated. It was filtered out and washed with 1M HCl solution and $Et_2O$ affording the desired alcohol in pure form.

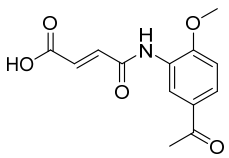
In the case of **61**, the reaction mixture was diluted with water and extracted with hexane. The aqueous phase was then brought to pH 3-4 by addition of a 10% citric acid solution. It was extracted with EtOAc three times, the combined organic phases were dried over $MgSO_4$ and evaporated under reduced pressure obtaining the final product **61** in pure form.




**115**, R = methoxy
**116**, R = methyl
**117**, R = fluoro
**118**, R = H

**40**, R = methoxy
**41**, R = methyl
**42**, R = fluoro
**43**, R = H

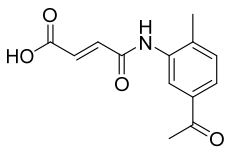
#### (*E*)-4-((5-Acetyl-2-methoxyphenyl)amino)-4-oxobut-2-enoic acid (40)



Yellow solid; Yield: 53 %; mp 172-174 °C; $^1$H NMR (500 MHz, MeOD): δ = 8.82 (s, 1H), 7.87 (d, $J$ = 8.3 Hz, 1H), 7.37 (d, $J$ = 15.4 Hz, 1H), 7.16 (d, $J$ = 8.6 Hz, 1H), 6.83 (d, $J$ = 15.6 Hz, 1H), 4.00 (s, 3H), 2.57 (s, 3H); $^{13}$C NMR (125 MHz, MeOD): δ = 199.2, 168.4, 164.6, 155.6, 138.0, 132.6, 131.0, 128.1, 127.8, 123.7, 111.5, 56.8, 26.5; IR (neat): ῡ = 3306, 3068, 2949, 2840, 1702, 1665,

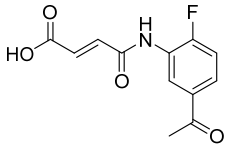
1589, 1542, 1516, 1445, 1422, 1359, 1302, 1269, 1213, 1174, 1134, 1023, 973, 886, 813 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{14}NO_5^+$, 264.0867; found, 264.0869.

**(*E*)-4-((5-Acetyl-2-methylphenyl)amino)-4-oxobut-2-enoic acid (41)**



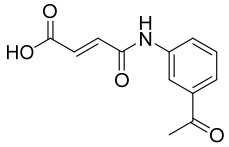
White solid; Yield: 60 %; mp 203-204 °C; $^1$H NMR (400 MHz, MeOD): δ = 8.14 (s, 1H), 7.80 (d, *J* = 8.1 Hz, 1H), 7.41 (d, *J* = 7.9 Hz, 1H), 7.28 (d, *J* = 15.3 Hz, 1H), 6.84 (d, *J* = 15.4 Hz, 1H), 6.76 (s, 1H), 2.58 (s, 3H), 2.35 (s, 3H). $^{13}$C NMR (125 MHz, DMSO-*d*$_6$): δ = 207.2, 176.3, 172.1, 147.1, 146.9, 146.0, 145.1, 140.9, 140.8, 135.4, 133.9, 36.6, 28.1; IR (neat): ῡ = 3266, 2925, 2604, 1714, 1659, 1607, 1573, 1536, 1417, 1288, 1259, 1231, 1177, 977, 912, 905, 822, 811, 706 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{13}NO_4^+$, 248.0917; found, 248.0916.

**(*E*)-4-((5-Acetyl-2-fluorophenyl)amino)-4-oxobut-2-enoic acid (42)**

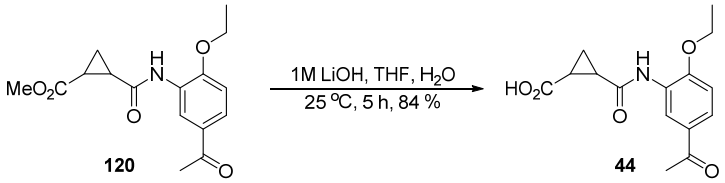


White solid; Yield: 60 %; mp 218-220 °C; $^1$H NMR (400 MHz, MeOD): δ = 8.76 (dd, *J* = 7.5, 2.0 Hz, 1H), 7.86 (ddd, *J* = 8.5, 4.7, 2.2 Hz, 1H), 7.32 (t, *J* = 9.5 Hz, 1H), 7.30 (d, *J* = 13.9 Hz, 1H), 6.86 (d, *J* = 15.4 Hz, 1H), 2.60 (s, 3H); $^{13}$C NMR (125 MHz, MeOD): δ = 198.5, 168.2, 164.8, 164.6, 159.2, 157.2, 155.9, 154.0, 140.8, 137.5, 137.2, 135.0, 133.2, 132.8, 127.9 (d, *J* = 9.0 Hz), 127.3 (d, *J* = 12.3 Hz), 126.3 (d, *J* = 11.6 Hz), 125.6 (d, *J* = 2.2 Hz), 125.5 (d, *J* = 7.8 Hz), 123.8, 116.8 (d, *J* = 20.8 Hz), 116.0 (d, *J* = 20.0 Hz), 102.4, 26.7, presence of rotamers; IR (neat): ῡ = 3446, 3297, 3084, 1707, 1666, 1611, 1601, 1546, 1487, 1417, 1362, 1335, 1293, 1267, 1192, 1112, 990, 925, 887, 825, 758, 721 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{11}FNO_4^+$, 252.0667; found, 252.0669.

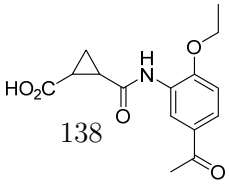
**(*E*)-4-((3-Acetylphenyl)amino)-4-oxobut-2-enoic acid (43)**



Off white solid; Yield: 90 %; mp 237-239 °C; $^1$H NMR (400 MHz, DMSO-*d*$_6$): δ = 13.03 (s, 1H), 10.72 (s, 1H), 8.26 (s, 1H), 7.92 (d, *J* = 7.9 Hz, 1H), 7.72 (d, *J* = 7.7 Hz, 1H), 7.51 (t, *J* = 7.9 Hz, 1H), 7.14 (d, *J* = 15.3 Hz, 1H), 6.69 (d, *J* = 15.4 Hz, 1H), 2.57 (s, 3H); $^{13}$C NMR (125 MHz, DMSO-*d*$_6$): δ = 197.6, 166.3, 161.9, 139.0, 137.4, 136.9, 131.1, 129.4, 124.1, 123.9, 118.6, 26.8; IR (neat): ῡ = 3349, 2835, 2682, 2570, 1703, 1661, 1609, 1552, 1490, 1421, 1363, 1341, 1307, 1281, 1166, 974, 879, 794 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{12}NO_4^+$, 234.0761; found, 234.0760.

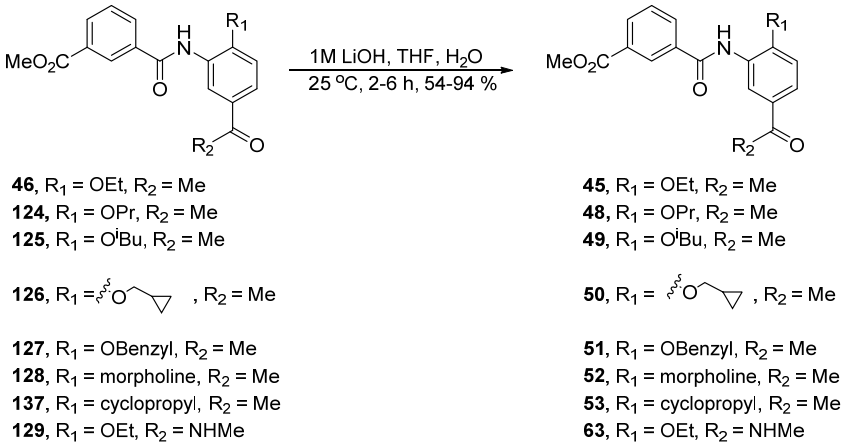



**2-((5-Acetyl-2-ethoxyphenyl)carbamoyl)cyclopropane-1-carboxylic acid (44)**



White solid; Yield: 84 %; mp 234-235 °C; $^1$H NMR (500 MHz, DMSO-*d*$_6$): δ = 12.50 (br, 1H), 9.65 (s, 1H), 8.53 (s, 1H), 7.73 (dd, *J* = 8.5, 1.7 Hz,

1H), 7.13 (d, $J$ = 8.6 Hz, 1H), 4.21 (q, $J$ = 6.9 Hz, 2H), 2.71 – 2.62 (m, 1H), 2.48 (s, 3H), 1.94 – 1.87 (m, 1H), 1.41 (t, $J$ = 7.0 Hz, 3H), 1.34 – 1.22 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 196.2, 173.1, 169.1, 152.7, 129.2, 127.1, 125.9, 121.9, 111.4, 64.3, 26.3, 23.3, 21.5, 14.4, 14.2; IR (neat): ῦ = 3408, 2992, 2414, 1716, 1642, 1598, 1575, 1517, 1425, 1366, 1320, 1271, 1191, 1130, 1082, 1037, 969, 930, 892, 879, 810 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{15}H_{17}NNaO_5^+$, 314.0999; found, 314.0996.




**46**, $R_1$ = OEt, $R_2$ = Me
**124,** $R_1$ = OPr, $R_2$ = Me
**125**, $R_1$ = O$^i$Bu, $R_2$ = Me

**126**, $R_1$ =  , $R_2$ = Me

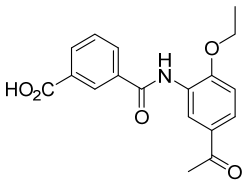
**127**, $R_1$ = OBenzyl, $R_2$ = Me
**128**, $R_1$ = morpholine, $R_2$ = Me
**137**, $R_1$ = cyclopropyl, $R_2$ = Me
**129**, $R_1$ = OEt, $R_2$ = NHMe

**45**, $R_1$ = OEt, $R_2$ = Me
**48**, $R_1$ = OPr, $R_2$ = Me
**49**, $R_1$ = O$^i$Bu, $R_2$ = Me

**50**, $R_1$ =  , $R_2$ = Me

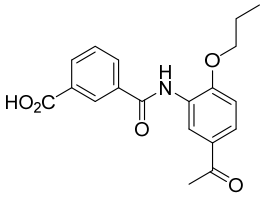
**51**, $R_1$ = OBenzyl, $R_2$ = Me
**52**, $R_1$ = morpholine, $R_2$ = Me
**53**, $R_1$ = cyclopropyl, $R_2$ = Me
**63**, $R_1$ = OEt, $R_2$ = NHMe

### 3-((5-Acetyl-2-ethoxyphenyl)carbamoyl)benzoic acid (45)



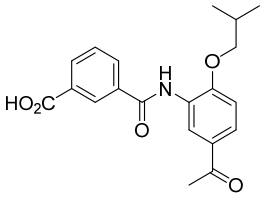

White solid; Yield: 60 %; mp 224-226 °C; $^1$H NMR (400 MHz, MeOD): δ = 8.65 (d, $J$ = 2.2 Hz, 1H), 8.61 – 8.59 (m, 1H), 8.25 (ddd, $J$ = 7.8, 1.6, 1.2 Hz, 1H), 8.17 (ddd, $J$ = 7.8, 1.8, 1.2 Hz, 1H), 7.90 (dd, $J$ = 8.7, 2.2 Hz, 1H), 7.67 (dd, $J$ = 8.0, 7.5 Hz, 1H), 7.17 (d, $J$ = 8.7 Hz, 1H), 4.27 (q, $J$ = 7.0 Hz, 2H), 2.59 (s, 3H), 1.50 (t, $J$ = 7.0 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 196.2, 166.8, 164.6, 155.2, 134.8, 132.3, 131.7, 131.2, 129.3, 129.0, 128.5, 127.4, 126.7, 124.7, 111.9, 64.4, 26.4, 14.4; IR (neat): ῦ = 3437, 3412, 2988, 1710, 1684, 1672, 1603, 1585, 1533, 1429, 1301, 1269, 1229, 1205, 1147, 1126, 1077, 1032, 904, 801, 732 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{18}H_{18}NO_5^+$, 328.1180; found, 328.1176.

### 3-((5-Acetyl-2-propoxyphenyl)carbamoyl)benzoic acid (48)



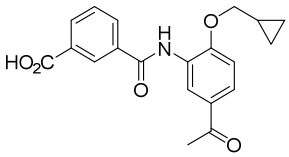

Beige solid; Yield: 77 %; mp 208-211 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 8.63 (dd, $J$ = 15.5, 7.4 Hz, 2H), 8.25 (d, $J$ = 7.7 Hz, 1H), 8.17 (d, $J$ = 7.7 Hz, 1H), 7.90 (d, $J$ = 8.5 Hz, 1H), 7.66 (t, $J$ = 7.7 Hz, 1H), 7.17 (d, $J$ = 8.6 Hz, 1H), 4.16 (t, $J$ = 6.3 Hz, 2H), 2.59 (s, 3H), 1.96 – 1.84 (m, 2H), 1.10 (t, $J$ = 7.4 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 189.7, 159.2, 157.9, 146.7, 126.8, 124.6, 123.3, 121.6, 120.7, 120.1, 119.0, 118.6, 115.3, 103.0, 62.3, 17.0, 14.0, 1.3, 1 C is missing due to overlapping; IR (neat): ῦ = 3427, 2965, 2940, 1719, 1665, 1639, 1590, 1507, 1434, 1404, 1389, 1333, 1262, 1206, 1158, 1070, 1035, 1009, 910, 810, 727 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{19}H_{20}NO_5^+$, 342.1336; found, 342.1341.

**3-((5-Acetyl-2-isobutoxyphenyl)carbamoyl)benzoic acid (49)**



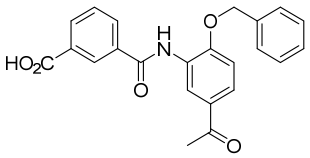
White solid; Yield: 94 %; mp 221-223 °C; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 13.21 (s, 1H), 9.79 (s, 1H), 8.52 (t, $J$ = 1.5 Hz, 1H), 8.26 (d, $J$ = 2.2 Hz, 1H), 8.20 – 8.16 (m, 1H), 8.16 – 8.12 (m, 1H), 7.86 (dd, $J$ = 8.6, 2.2 Hz, 1H), 7.68 (t, $J$ = 7.7 Hz, 1H), 7.20 (d, $J$ = 8.7 Hz, 1H), 3.92 (d, $J$ = 6.4 Hz, 2H), 2.54 (s, 3H), 2.06 (dt, $J$ = 13.2, 4.8 Hz, 1H), 0.99 (s, 3H), 0.97 (s, 3H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ = 196.2, 166.7, 164.5, 155.7, 134.8, 132.2, 131.6, 131.2, 129.3, 129.0, 128.3, 127.6, 126.7, 124.9, 111.8, 74.5, 27.7, 26.4, 18.9; IR (neat): ῦ = 3375, 2963, 2933, 2877, 2469, 1682, 1644, 1596, 1581, 1510, 1426, 1411, 1356, 1316, 1297, 1270, 1214, 1128, 1107, 1017, 956, 885, 810, 724, 724 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{20}H_{21}NNaO_4^+$, 378.1312; found, 378.1309.

**3-((5-Acetyl-2-(cyclopropylmethoxy)phenyl)carbamoyl)benzoic acid (50)**



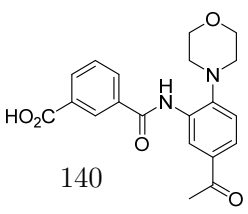
Beige solid; Yield: 54 %; mp 221-223 °C; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 13.23 (s, 1H), 9.76 (s, 1H), 8.53 (s, 1H), 8.32 (d, $J$ = 2.2 Hz, 1H), 8.21 – 8.17 (m, 1H), 8.17 – 8.13 (m, 1H), 7.84 (dd, $J$ = 8.6, 2.3 Hz, 1H), 7.69 (t, $J$ = 7.8 Hz, 1H), 7.20 (d, $J$ = 8.7 Hz, 1H), 4.03 (d, $J$ = 6.8 Hz, 2H), 2.54 (s, 3H), 1.33 – 1.20 (m, 1H), 0.60 – 0.51 (m, 2H), 0.42 – 0.34 (m, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 196.2, 166.7, 164.4, 155.2, 134.8, 132.3, 131.7, 131.2, 129.4, 129.1, 128.3, 127.4, 126.8, 124.4, 112.3, 72.9, 26.4, 9.9, 3.0; IR (neat): ῦ = 3075, 2957, 2927, 1716, 1670, 1597, 1582, 1510, 1432, 1412, 1396, 1291, 1276, 1262, 1217, 1160, 1111, 1069, 1023, 991, 906, 796, 728 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{20}H_{19}NNaO_5^+$, 376.1155; found, 376.1151.

**3-((5-Acetyl-2-(benzyloxy)phenyl)carbamoyl)benzoic acid (51)**



White solid; Yield: 79 %; mp 210-215 °C; $^{1}$H NMR (400 MHz, DMSO-$d_6$): δ = 9.88 (s, 1H), 8.54 (s, 1H), 8.30 (d, $J$ = 1.9 Hz, 1H), 8.12 (t, $J$ = 8.1 Hz, 2H), 7.85 (dd, $J$ = 8.6, 2.1 Hz, 1H), 7.61 (t, $J$ = 7.7 Hz, 1H), 7.52 (d, $J$ = 7.2 Hz, 2H), 7.37 (t, $J$ = 7.4 Hz, 2H), 7.30 (d, $J$ = 7.4 Hz, 1H), 7.26 (d, $J$ = 8.7 Hz, 1H), 5.32 (s, 2H), 2.54 (s, 3H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ = 196.2, 167.1, 165.0, 155.0, 136.5, 134.5, 132.2, 130.7, 129.7, 128.4, 128.3, 127.8, 127.3, 127.1, 125.1, 112.5, 69.9, 26.4, three carbons are missing due to overlapping; IR (neat): ῦ = 3414, 3085, 1713, 1660, 1637, 1593, 1584, 1510, 1434, 1400, 1390, 1323, 1296, 1274, 1221, 1165, 1021, 995, 896, 803, 724, 702 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{23}H_{20}NO_5^+$, 390.1336; found, 390.1335.

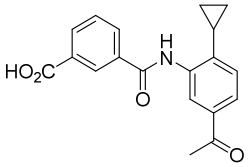
**3-((5-Acetyl-2-morpholinophenyl)carbamoyl)benzoic acid (52)**



White solid; Yield: 76 %; mp 249-252 °C; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 13.32 (s, 1H), 9.93 (s, 1H), 8.53 (s, 1H), 8.43 (d, $J$ = 1.7 Hz, 1H), 8.21 (d, $J$ = 7.7 Hz, 1H), 8.16 (d, $J$ = 7.7 Hz, 1H), 7.82 (dd, $J$ = 8.4, 1.4

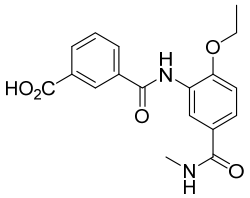
Hz, 1H), 7.71 (t, $J$ = 7.7 Hz, 1H), 7.29 (d, $J$ = 8.4 Hz, 1H), 3.78 (t, $J$ = 3.9 Hz, 4H), 2.99 (t, $J$ = 4.1 Hz, 4H), 2.55 (s, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ = 196.7, 166.8, 164.2, 149.3, 134.7, 132.5, 131.9, 131.8, 131.3, 131.2, 129.3, 128.2, 126.5, 124.2, 119.7, 66.4, 51.0, 26.6; IR (neat): ῡ = 3346, 2981, 2831, 1723, 1676, 1644, 1598, 1577, 1529, 1457, 1432, 1283, 1234, 1215, 1112, 1069, 935, 918, 894, 848, 824, 726 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{20}H_{21}N_2O_5^+$, 369.1445; found, 369.1444.

### 3-((5-Acetyl-2-cyclopropylphenyl)carbamoyl)benzoic acid (53)

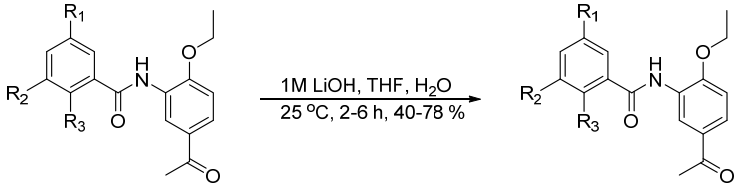


Pale yellow solid; Yield: 57 %; mp 95-100 °C; $^1$H NMR (400 MHz, DMSO-$d_6$): δ = 10.29 (s, 1H), 8.58 (s, 1H), 8.21 – 8.09 (m, 3H), 7.94 (s, 1H), 7.77 (d, $J$ = 8.4 Hz, 1H), 7.61 (t, $J$ = 7.6 Hz, 1H), 7.10 (d, $J$ = 8.6 Hz, 1H), 2.55 (s, 3H), 2.18 – 2.04 (m, 1H), 1.06 – 0.93 (m, 2H), 0.76 – 0.69 (m, 2H);$^{13}$C NMR (100 MHz, DMSO-$d_6$): δ = 197.0, 167.1, 165.4, 145.0, 137.1, 134.5, 134.4, 132.2, 130.9, 128.5, 128.5, 126.3, 126.2, 124.9, 26.6, 11.5, 9.0, one C is missing due to overlapping; IR (neat): ῡ = 3282, 2920, 1683, 1638, 1608, 1571, 1532, 1416, 1358, 1294, 1260, 1231, 1101, 1077, 1046, 1022, 882, 819, 797, 728 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{19}H_{17}NNaO_4^+$, 346.1050; found, 346.1047.

### 3-((2-Ethoxy-5-(methylcarbamoyl)phenyl)carbamoyl)benzoic acid (63)



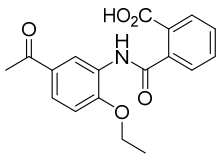
Pale brown solid; Yield: 62 %; mp 226-229 °C; $^1$H NMR (400 MHz, MeOD): δ = 8.60 (t, $J$ = 1.5 Hz, 1H), 8.42 (d, $J$ = 2.2 Hz, 1H), 8.27 – 8.22 (m, 1H), 8.19 – 8.15 (m, 1H), 7.73 – 7.60 (m, 2H), 7.14 (d, $J$ = 8.7 Hz, 1H), 4.23 (q, $J$ = 6.9 Hz, 2H), 2.92 (s, 3H), 1.48 (t, $J$ = 7.0 Hz, 3H); $^{13}$C NMR (100 MHz, MeOD): δ = 170.3, 168.8, 167.4, 154.6, 136.3, 134.0, 132.9, 132.8, 132.7, 130.2, 130.1, 130.0, 129.6, 127.8, 127.7, 126.7, 123.7, 112.6, 65.8, 27.0, 15.0, presence of rotamers; IR (neat): ῡ = 3398, 3081, 2981, 2525, 1689, 1661, 1601, 1573, 1467, 1429, 1395, 1336, 1267, 1234, 1195, 1167, 1123, 1029, 913, 782, 760, 735 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{18}H_{18}N_2O_5^+$, 343.1289; found, 343.1283.




**131**, $R_1$ = H, $R_2$ = H, $R_3$ = $CO_2Me$
**138**, $R_1$ = pyridine, $R_2$ = $CO_2Me$, $R_3$ = H
**139**, $R_1$ = furane, $R_2$ = $CO_2Me$, $R_3$ = H
**133**, $R_1$ = $CH_2OMe$, $R_2$ = $CO_2Me$, $R_3$ = H
**134**, $R_1$ = $CH_2OTHP$, $R_2$ = $CO_2Me$, $R_3$ = H

**47**, $R_1$ = H, $R_2$ = H, $R_3$ = $CO_2H$
**57**, $R_1$ = pyridine, $R_2$ = $CO_2H$, $R_3$ = H
**58**, $R_1$ = furane, $R_2$ = $CO_2H$, $R_3$ = H
**60**, $R_1$ = $CH_2OMe$, $R_2$ = $CO_2H$, $R_3$ = H
**61**, $R_1$ = $CH_2OTHP$, $R_2$ = $CO_2H$, $R_3$ = H

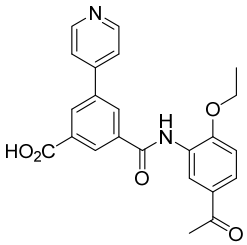
### 2-((5-Acetyl-2-ethoxyphenyl)carbamoyl)benzoic acid (47)



White solid; Yield: 57 %; mp 143-146 °C; $^1$H NMR (400 MHz, MeOD): δ = 8.75 (s, 1H), 8.02 (d, $J$ = 7.6 Hz, 1H), 7.86 (d, $J$ = 8.9 Hz, 1H), 7.71 – 7.62

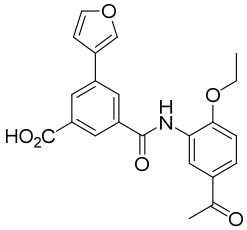
(m, 1H), 7.63 – 7.55 (m, 2H), 7.13 (d, $J$ = 8.4 Hz, 1H), 4.22 (q, $J$ = 6.9 Hz, 2H), 2.59 (s, 3H), 2.59 (s, 3H); $^{13}$C NMR (126 MHz, MeOD): δ = 198.7, 197.9, 169.8, 169.1, 168.0, 154.3, 151.3, 138.4, 136.6, 132.7, 131.8, 130.9, 130.6, 130.0, 129.5, 129.5, 128.8, 128.1, 127.4, 126.6, 123.4, 120.4, 114.1, 110.9, 109.8, 64.5, 63.8, 48.2, 48.1, 48.1, 47.9, 47.9, 47.8, 47.7, 47.7, 47.7, 47.7, 47.6, 47.4, 47.3, 47.3, 47.1, 25.1, 24.9, 13.7, 13.4, presence of rotamers; IR (neat): ῦ = 3272, 2970, 2936, 1713, 1662, 1603, 1583, 1540, 1498, 1436, 1275, 1254, 1220, 1159, 1128, 1043, 933, 884, 809, 786, 730, 710, 704 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{18}H_{18}NO_5^+$, 328.1180; found, 328.1181.

**3-((5-Acetyl-2-ethoxyphenyl)carbamoyl)-5-(pyridin-4-yl)benzoic acid (57)**



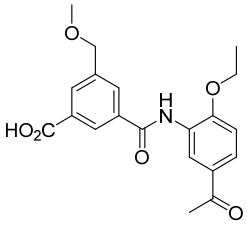
White solid; Yield: 40 % over two steps; mp 270-276 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 9.91 (s, 1H), 8.70 (d, $J$ = 5.7 Hz, 2H), 8.55 (s, 1H), 8.45 (s, 1H), 8.42 (s, 1H), 8.30 (s, 1H), 7.87 (dd, $J$ = 8.5, 2.0 Hz, 1H), 7.83 (d, $J$ = 4.9 Hz, 2H), 7.22 (d, $J$ = 8.7 Hz, 1H), 4.21 (q, $J$ = 6.9 Hz, 2H), 2.55 (s, 3H), 1.37 (t, $J$ = 6.9 Hz, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ= 196.2, 166.3, 164.1, 155.6, 145.4, 136.3, 136.0, 132.5, 131.0, 130.8, 130.4, 129.4, 127.8, 126.4, 125.5, 123.6, 112.0, 64.4, 26.4, 14.4, 1 C is missing due to overlapping; IR (neat): ῦ = 3441, 2969, 1682, 1591, 1540, 1432, 1357, 1331, 1260, 1225, 1203, 1066, 1027, 797, 748, 732, 707 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{23}H_{21}N_2O_5^+$, 405.1445; found, 405.1440.

**3-((5-Acetyl-2-ethoxyphenyl)carbamoyl)-5-(furan-3-yl)benzoic acid (58)**



Beige solid; Yield: 55 %; mp 210-214 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 13.37 (s, 1H), 9.90 (s, 1H), 8.42 (s, 1H), 8.37 (s, 1H), 8.33 (s, 1H), 8.29 (s, 1H), 7.87 (dd, $J$ = 8.5, 1.9 Hz, 1H), 7.83 (s, 1H), 7.22 (d, $J$ = 8.4 Hz, 1H), 7.13 (s, 1H), 4.21 (q, $J$ = 6.6 Hz, 2H), 2.55 (s, 3H), 1.38 (t, $J$ = 6.8 Hz, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ = 196.3, 166.8, 164.5, 155.4, 144.8, 140.6, 135.5, 132.9, 131.9, 129.3, 128.9, 128.6, 127.6, 127.0, 126.6, 125.2, 124.7, 111.9, 108.7, 64.4, 26.5, 14.5; IR (neat): ῦ = 3424, 3147, 2988, 2935, 1719, 1678, 1658, 1588, 1540, 1492, 1432, 1361, 1336, 1266, 1202, 1183, 1141, 1070, 1024, 907, 890, 873, 808, 795, 741 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{22}H_{19}NO_6^+$, 394.1285; found, 394.1279.

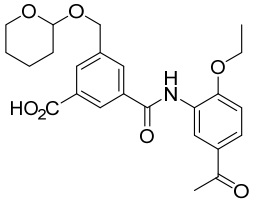
**3-((5-Acetyl-2-ethoxyphenyl)carbamoyl)-5-(methoxymethyl)benzoic acid (60)**



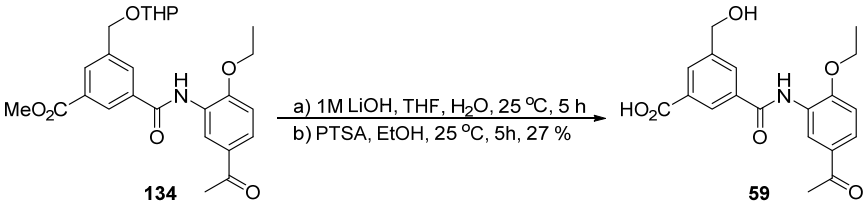
White solid; Yield: 50 %; mp 185-187 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 13.29 (br, 1H), 9.86 (s, 1H), 8.44 (s, 1H), 8.26 (s, 1H), 8.13 (s, 1H), 8.09 (s, 1H), 7.86 (d, $J$ = 8.7 Hz, 1H), 7.21 (d, $J$ = 8.6 Hz, 1H), 4.57 (s, 2H), 4.20 (q, $J$ = 6.7 Hz, 2H), 3.36 (s, 3H), 2.54 (s, 3H), 1.37 (t, $J$ = 6.8 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 196.3, 166.8, 164.6, 155.4, 139.6, 134.9, 131.2, 131.0, 130.6, 129.4, 127.6, 126.7, 125.0, 111.9, 72.7, 64.4, 57.9, 26.5, 14.5;

one C is missing due to overlapping; IR (neat): $\tilde{\upsilon}$ = 3443, 2986, 2922, 1683, 1592, 1539, 1432, 1361, 1263, 1243, 1190, 1135, 1113, 1035, 893, 809, 742, 708 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{20}H_{21}NNaO_6^+$, 394.1261; found, 394.1263.

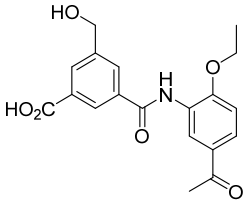
### 3-((5-Acetyl-2-ethoxyphenyl)carbamoyl)-5-(((tetrahydro-2*H*-pyran-2-yl)oxy)methyl)benzoic acid (61)



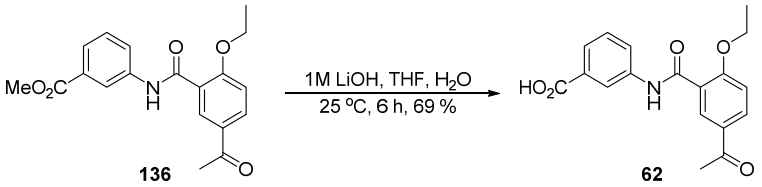
White solid; Yield: 78 %; mp 165-170 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 13.23 (br, 1H), 9.80 (s, 1H), 8.44 (s, 1H), 8.31 (d, *J* = 1.9 Hz, 1H), 8.14 (s, 1H), 8.11 (s, 1H), 7.85 (dd, *J* = 8.6, 2.1 Hz, 1H), 7.20 (d, *J* = 8.7 Hz, 1H), 4.82 (d, *J* = 12.6 Hz, 1H), 4.75 (t, *J* = 3.3 Hz, 1H), 4.62 (d, *J* = 12.6 Hz, 1H), 4.21 (q, *J* = 6.9 Hz, 2H), 3.81 (ddd, *J* = 11.1, 8.3, 2.9 Hz, 1H), 3.53 – 3.49 (m, 1H), 2.54 (s, 3H), 1.84 – 1.65 (m, 2H), 1.62 – 1.44 (m, 4H), 1.37 (t, *J* = 6.9 Hz, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ = 196.4, 166.8, 164.7, 155.3, 139.8, 135.1, 131.4, 131.2, 130.8, 129.5, 127.6, 127.5, 126.9, 124.8, 112.0, 97.8, 67.6, 64.5, 61.6, 30.3, 26.5, 25.1, 19.2, 14.5; IR (neat): $\tilde{\upsilon}$ = 3427, 2938, 2873, 1718, 1679, 1647, 1590, 1536, 1436, 1343, 1296, 1260, 12001, 1185, 1121, 1073, 1027, 976, 957, 899, 809, 743 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{24}H_{27}NNaO_7^+$, 464.1680; found, 464.1678.



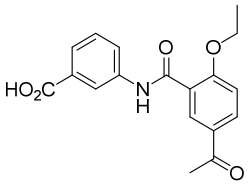

### 3-((5-Acetyl-2-ethoxyphenyl)carbamoyl)-5-(hydroxymethyl)benzoic acid (59)



White solid; Yield: 27 % over two steps; mp 225-228 °C; $^1$H NMR (500 MHz, DMSO-$d_6$): δ = 13.24 (s, 1H), 9.83 (s, 1H), 8.39 (s, 1H), 8.27 (s, 1H), 8.12 (d, *J* = 4.9 Hz, 2H), 7.86 (d, *J* = 8.5 Hz, 1H), 7.21 (d, *J* = 8.6 Hz, 1H), 5.49 (t, *J* = 5.6 Hz, 1H), 4.65 (d, *J* = 5.5 Hz, 2H), 4.20 (q, *J* = 6.9 Hz, 2H), 2.54 (s, 3H), 1.37 (t, *J* = 6.9 Hz, 3H); $^{13}$C NMR (126 MHz, DMSO-$d_6$): δ = 196.3, 167.0, 164.7, 155.3, 143.8, 134.7, 131.0, 130.1, 129.8, 129.4, 127.5, 126.9, 126.8, 124.9, 111.9, 64.4, 62.2, 26.5, 14.5; IR (neat): $\tilde{\upsilon}$ = 3391, 3262, 2975, 2933, 1719, 1646, 1604, 1581, 1541, 1498, 1477, 1437, 1361, 1336, 1298, 1274, 1232, 1206, 1183, 1147, 1054, 1037, 987, 904, 893, 821, 796, 753, 710 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{19}H_{20}NO_6^+$, 358.1285; found, 358.1285.




**3-(5-Acetyl-2-ethoxybenzamido)benzoic acid (62)**



Pale yellow solid; Yield: 69 %; mp 209-212 °C; $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ = 12.96 (br, 1H), 10.34 (s, 1H), 8.39 (s, 1H), 8.20 (d, $J$ = 1.9 Hz, 1H), 8.09 (dd, $J$ = 8.7, 2.0 Hz, 1H), 7.91 (d, $J$ = 7.8 Hz, 1H), 7.68 (d, $J$ = 7.7 Hz, 1H), 7.48 (t, $J$ = 7.9 Hz, 1H), 7.28 (d, $J$ = 8.7 Hz, 1H), 4.26 (q, $J$ = 6.9 Hz, 2H), 2.57 (s, 3H), 1.40 (t, $J$ = 6.9 Hz, 3H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ = 196.0, 167.1, 164.0, 159.5, 139.1, 132.4, 131.4, 130.2, 129.4, 129.0, 124.9, 124.4, 123.7, 120.3, 112.7, 64.7, 26.5, 14.4; IR (neat): $\tilde{\upsilon}$ = 3343, 2981, 1719, 1670, 1596, 1552, 1488, 1409, 1362, 1299, 1265, 1203, 1149, 1110, 1080, 1025, 921, 890, 811, 756 cm$^{-1}$; HRMS (ESI), $m/z$: calcd for $C_{18}H_{17}NNaO_5^+$, 350.0999; found, 350.0995.

### VI. Synthesis of novel bromodomain ligands based on docking studies (64-66)

Due to an on-going patent application, the synthesis of compounds **64-66** will not be disclosed.

## 3.3.5 Bromodomain expression and purification

Proteins were purified as described previously.[31] Briefly, His-tagged bromodomains were expressed in *Escherichia coli* BL21(DE3) cells upon induction with isopropyl thio-beta-D-galactoside (IPTG, final concentration 0.1 mM) for 16 h at 18 °C. Bacteria were lysated and (when required) the resulting extract was treated to remove DNA, either passing it through a DEAE column or adding 0.15% polyethylenimine (PEI). The His-tagged proteins were purified on HisTrap columns (GE Healthcare) and eluted using a step gradient of imidazole. The poly-Histidine tags were removed by overnight incubation with His-tagged tobacco etch virus (TEV) protease purified in-house (if required by the purification protocol, in the meantime the sample was exchanged via dialysis). A size-exclusion chromatography step (HiLoad 16/600 Superdex75 column) and a Ni-affinity chromatography step were subsequently performed to finally purify the cleaved bromodomains. Samples were then concentrated, flash frozen and stored at –80 °C.

### 3.3.6 X-ray crystallography

#### I. Crystallization, Data Collection, and Structure Determination

Crystals of the CREBBP bromodomain were grown at 4°C using the hanging drop vapor diffusion method. A 50 mM solution of compound **45** (in 100 % DMSO) was added into the CREBBP protein to reach a final DMSO concentration of 1 % (v/v) and the mixture was incubated on ice for 1 hour before crystallization. Then equal volumes of protein (with compound **45**) and reservoir solutions (0.1 M MES pH 6.5, 0.10 $MgCl_2$, 20 % PEG 6000, 10 % ethylene glycol) were mixed and crystals appeared after 1 to 2 days. The crystals were flash-frozen in liquid nitrogen with extra 10% ethylene glycol as cryoprotectant for measurements. Data sets were collected on a PILATUS 6MF detector at the Swiss Light Source beamline X06SA of the Paul Scherrer Institute (Villigen, Switzerland) and indexed, integrated and scaled with the XDS[174] and CCP4 programs.[175] The structures were solved by molecular replacement with PHASER[176] using the CREBBP structure (PDB entry 4NR5) as a search model and refined with PHENIX.[177] The atomic coordinates and structure factors of CREBBP in complex with inhibitor **45** have been deposited with the Protein Data Bank as entry 4TQN.

**Table 7**

| | Compound **45** |
|---|---|
| Space group | P1 21 1 |
| Unit cell | |
| a (Å) | 24.94 |
| b (Å) | 42.94 |
| c (Å) | 51.98 |
| alpha | 90.00 |
| beta | 97.24 |
| gamma | 90.00 |
| Resolution range (Å) | 42.94 -1.70 |
| Unique reflections | 12119(1768) |
| $<I/\sigma(I)>$ | 15.6(5.3) |
| R merge | 0.068(0.380) |
| Completeness (%) | 99.9(99.2) |
| Multiplicity | 6.5(6.2) |
| Refinement | |
| Resolution range (Å) | 33.00-1.70 |
| R factor/R free | 0.1813/0.1990 |
| Mean B factors (A2) | 23.24 |
| RMS bonds (Å) | 0.006 |
| RMS angles (°) | 1.155 |

**II. Composite Omit Map of Ligand 45**

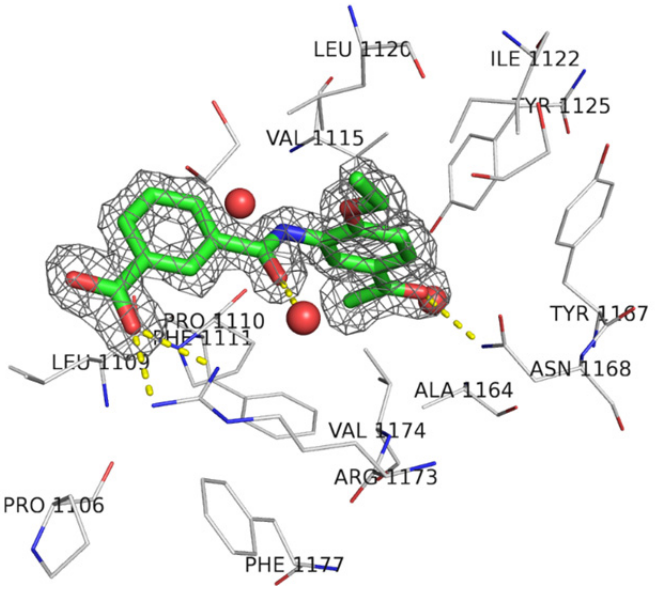



**Figure 27.** $2mF_o$ – $DF_c$ electron density maps contoured at 1σ (grey mesh) were generated in a region within 1.6 Å for compound **45** using PHENIX and Pymol.

### 3.3.7 Thermal shift measurements

Thermal shift measurements were carried out as previously described[17] with a final volume of 20 µl, ligand and protein concentrations 100 µM (50 µM for compound **57**) and 2 µM, respectively. The reported values ($\Delta T_m$) are calculated as the difference between the transition midpoints of an individual sample and the average of the reference wells (containing the protein and the DMSO only) in the same plate. DMSO concentration was kept at 0.2% (v/v).

**I. Phylogenetic tree**

The bromodomain sequence aligment previously reported by Filippakopoulos et al. was used excluding the residues from the plasmid.[31] The evolutionary history was inferred using the Neighbor-Joining method.[178] The optimal tree with the sum of branch length = 21.33784094 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method[179] and are in the units of the number of amino acid substitutions per site. The analysis involved 61 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 57 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.[180]

**II. Selectivity profiles**

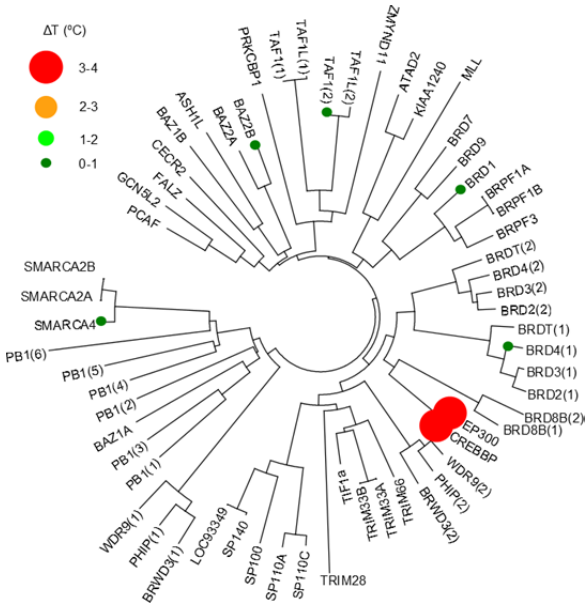



**Figure 28.** Bromodomain phylogenetic tree showing the selectivity evaluation of **45** as determined by thermal shift measurements ($\Delta T_m$) against a panel of different bromodomains. Sphere size and color indicate relative $\Delta T_m$ values according to the legend.

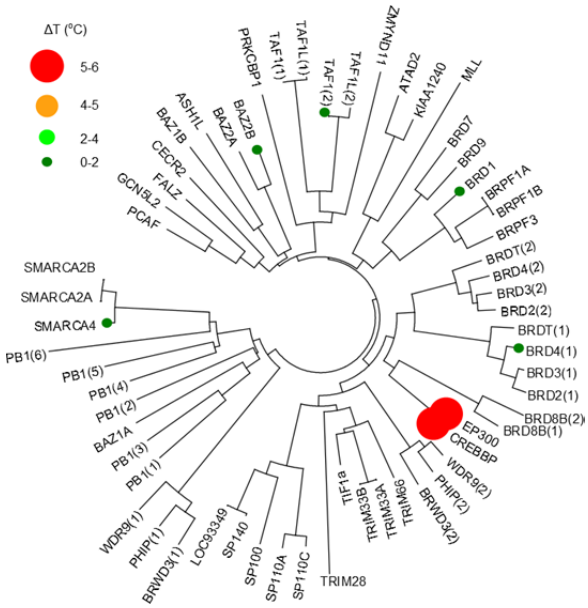



**Figure 29.** Bromodomain phylogenetic tree showing the selectivity evaluation of compounds **58**, **60**, and **61** as determined by thermal shift measurements ($\Delta T_m$) against a panel of different bromodomains. Sphere size and color indicate relative $\Delta T_m$ values according to the legend, which is different from the one of the previous figure.

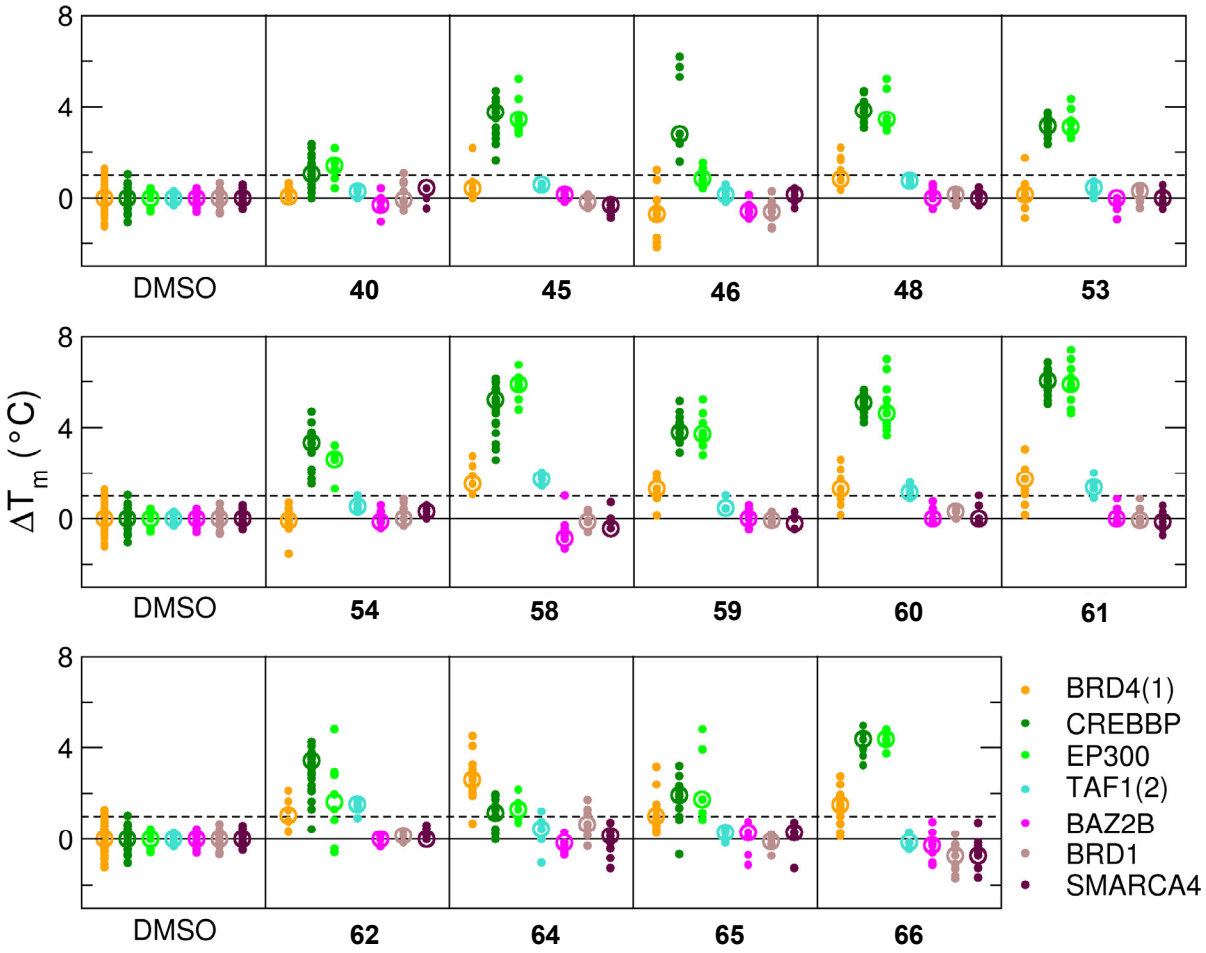



**Figure 30.** Selectivity evaluation of compounds **40**, **45**, **46**, **48**, **53**, **54**, **58-62**, **64-66** as determined by thermal shift measurements ($\Delta T_m$) against a panel of seven bromodomains. Independent measurements are shown as dots and the median is shown as a circle. The dashed line at 1°C is an arbitrary threshold based on the scatter plot in Figure 31.

### 3.3.8 TR-FRET assays

TR-FRET assays were carried out in duplicate at BPS Bioscience using a recombinant CBP bromodomain (BPS catalogue #31128) and the BET Ligand (BPS catalogue #33000) as provided in the CREBBP TR-FRET Assay Kit (BPS catalogue #32619). A 10 mM solution of the compound under investigation in DMSO was prepared and shipped to BPS Bioscience, where it was tested at 10 concentrations over the range of 0.001-10 µM (compounds **45**, **46**, **48-52**, **54-58**, **60-62**) or 0.01-100 µM (compounds **40-44**, **47**, **53**, **59**, **63**) . Each compound solution was then diluted in water to obtain a 10% DMSO solution. 2 µL of this dilution were added to a 20 µL reaction mixture (12.5 nM CBP, 125 nM BET Ligand, including FRET dyes and the amount of compound needed to reach the indicated concentration in the Table below). The resulting mixture was incubated for 2 hours at room temperature prior to reading the TR

(time resolved)-FRET signal using a Tecan Infinite M1000 plate reader. The negative control consisted of the aforementioned mixture in which the buffer was added in place of compound. TR-FRET were recorded as the ratio of the fluorescence of the acceptor and the donor dyes (acceptor/donor).

The TR-FRET data was analyzed using Graphpad Prism software. The percent activity in the presence of each compound was calculated according to the following equation: % activity = [(F- Fb)/(Ft – Fb)]×100, where Ft is the TR-FRET signal in the absence of any compound (100 % activity), Fb the TR-FRET signal in the absence of the bromodomain (0 % activity) and F the TR-FRET signal in the presence of the compound. The percent inhibition was calculated according to the following equation: % inhibition = 100 - % activity. The values of % activity versus a series of compound concentrations were then plotted using non-linear regression analysis of Sigmoidal dose-response curve generated with the equation Y=B+(T-B)/1+$10^{((\text{LogIC50-X})\times\text{Hill Slope})}$, where Y=percent activity, B=minimum percent activity, T=maximum percent activity, X= logarithm of compound and Hill Slope=slope factor or Hill coefficient. The $IC_{50}$ value corresponds to the concentration causing a half-maximal percent activity.

### 3.3.9 BROMOscan assays

$K_d$ determinations by means of BROMOscan technology was carried out at DiscoveRx. E. *coli* derived from BL21 strain was used as host to grow T7 phage strains displaying the bromodomains. E. *coli*, grown to log-phase, were infected with T7 phage (from a frozen stock, being the multiplicity of infection 0.4) and incubated while shaking at 32 °C for 90-150 minutes, until lysis. In order to remove cell debris, lysates were centrifuged at 5,000 x g and filtered (0.2 µm). Affinity resins were obtained by treating streptavidin-coated magnetic beads with biotinylated acetylated peptide ligands for 30 minutes at 25°C. Those beads were then blocked with excess of biotin and washed with blocking buffer (SeaBlock (Pierce), 1 % bovine serum albumin, BSA, 0.05 % Tween20, 1 mM dithiothreitol, (DTT) removing the unbound ligand and reducing non-specific phage binding.

During the experiment, the bromodomain, ligand-bound affinity beads and test compounds were combined in a buffer composed of 17% SeaBlock, 0.33x phosphate-buffered solution, PBS, 0.04% Tween20, 0.02% BSA, 0.004% sodium azide and 7.4 mM DTT. Test compounds were prepared as 50 mM in pure DMSO and diluted to 5 mM with monoethylene gycol, MEG (100× concentrated in respect to the top screening concentration, 50 µM). During the assay a DMSO and MEG final concentration of 0.1% and 0.9% respectively was used. The assays were carried out in polystyrene 96-well plates in a final volume of 0.135 mL. The assay plates were incubated at 25 °C with shaking for 1 hour and the affinity beads were washed with a buffer composed of 0.05% Tween 20 in PBS. The beads were then re-suspended in the elution buffer (1x PBS, 0.05% Tween 20, 2 µM non-biotinylated affinity ligand) and incubated at 25°C with shaking for 30 minutes. The bromodomain concentration in the eluates was measured by

qPCR. Binding constants ($K_d$) were calculated with a standard dose-response curve using the Hill equation and curves were fitted using a non-linear least square fit with the Levenberg-Marquardt algorithm.

### 3.3.10 ITC Experiments

Isothermal Titration Calorimetry experiments were performed on a VP-ITC instrument (MicroCal, Inc., Northampton, MA). Protein samples thoroughly dialyzed against the same batch of buffer in order to minimize artifacts due to minor differences in buffer composition and the protein concentration was determined after a filtering through a 0.22 µm pore-size filter.

Bromodomains (300-500 µM) were injected into the 1.4-mL sample cell containing the compound (50 µM) dissolved into the ITC buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1% DMSO). The titration experiments were carried out at 15°C while stirring at 300 rpm: after a control injection of 2 µL, 29 10-µL injections (10 s duration, with a 4 min interval between) were performed. The raw data were integrated, normalized for concentration, and analyzed using a single binding site model, supplied with the MicroCal Origin software package to obtain the apparent $K_d$ values ($K_d^{app}$). The actual $K_d$ values were calculated assuming that DMSO acted as a competing ligand[131] according to the equation:[181]

$$K_d = \frac{K_d^{app}}{1 + \frac{[DMSO]}{K_d^{DMSO}}}$$

where [DMSO] and $K_d^{DMSO}$ indicate the DMSO concentration in the ITC buffer and the dissociation constant of the DMSO for the bromodomain under investigation, respectively (Figure S9).

The $K_d^{DMSO}$ value for CREBBP was calculated titrating the histone H3K56ac peptide (ac-IRRYQ(Kac)STELLY-am, where ac-, -am and Kac, indicate acetylation, amidation and acetylated lysine side chain, respectively, purchased at GenScript) into a CREBBP bromodomain solution in the presence and absence of 0.75% DMSO with protein and peptide concentrations of 688 or 740 µM and 85 µM, respectively (Figure 35).

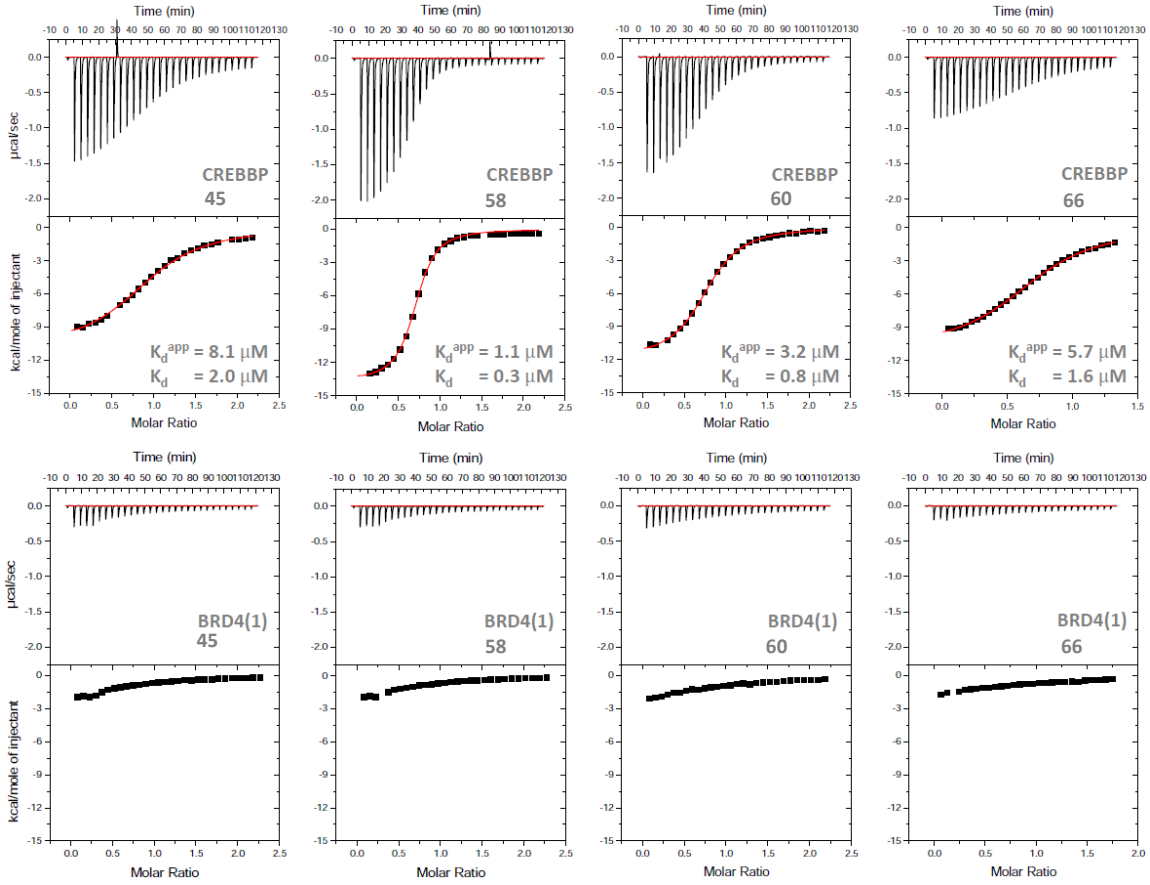
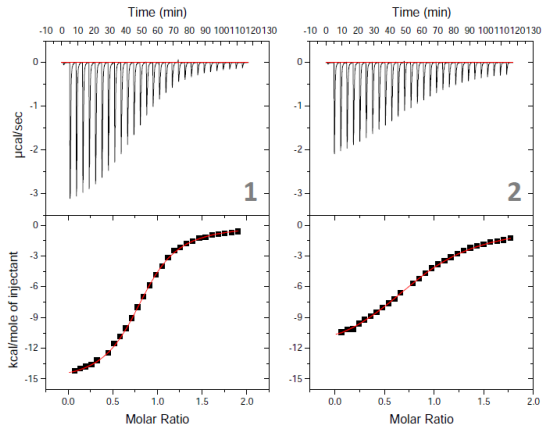



**Figure 31**. ITC titration curves for the binding of the CREBBP and BRD4(1) bromodomains (upper and lower panels, respectively) to compounds **45**, **58**, **60** and **66**.




| exp | [DMSO] (M) | app $K_d$ (µM) |
|---|---|---|
| 1 | 0.000 | 4.77 |
| 2 | 0.106 | 15.71 |

$$K_{d_{cmpd}} = \frac{K_{d_{cmpd}}^{app}}{1 + \frac{[DMSO]}{K_{d_{DMSO}}}}$$

$$K_{d_{DMSO}} = 46.3\ mM$$

**Figure 32**. The N-terminally acetylated, C-terminaly amidated IRRYQ(Kac)STELLY peptide was titrated into a CREBBP bromodomain solution in the absence of DMSO (experiment 1) and in the presence of 0.75% DMSO (experiment 2). The $K_d$ for DMSO was calculated as shown in the right panel.

**Table 8.** Thermodynamic parameters measured by ITC for the CREBBP bromodomain. ΔG and ΔH values are given in kcal/mol, ΔS values are given in cal/mol/T.

| Cmpd | **N** | **$K_d^{app}$ (μM)** | **ΔG** | **ΔH** | **ΔS** |
|---|---|---|---|---|---|
| **45** | 0.99 ± 0.01 | 8.13 ± 0.43 | –6.7 | –10.9 ± 0.1 | –14.5 |
| **58** | 0.70 ± 0.00 | 1.11 ± 0.08 | –7.9 | –13.7 ± 0.1 | –20.1 |
| **60** | 0.79 ± 0.00 | 3.25 ± 0.10 | –7.2 | –11.9 ± 0.0 | –16.2 |
| **66** | 0.75 ± 0.00 | 5.69 ± 0.16 | –6.9 | –10.8 ± 0.0 | –13.6 |

### 3.3.11 Cell culture and cytotoxicity measurements

MDA-MB-231, HT-29 cells (obtained from the UZH Cancer Institute) and HeLa cells (obtained from Dr. Nathan Luedtke, Chemistry Department, UZH) were cultured in DMEM supplemented with 10 % (v/v) fetal bovine serum. K562 cells (obtained from Dr. Silvio Hemmi, Institute of Molecular Life Sciences, UZH), HL-60, ML2 (obtained from Dr. Nathan Luedtke, Chemistry Department, UZH) and HOP-92 (purchased from the NCI) were cultured using RPMI medium supplemented with 10 % (v/v) fetal bovine serum. Finally, AML3, PL-21 and MOLM-13 (obtained from Dr. Nathan Luedtke, Chemistry Department, UZH) were cultured using RPMI medium supplemented with 20 % (v/v) fetal bovine serum. All the media were additionally supplemented with 100 units/mL of penicillin, 100 μg/mL of streptomycin, 4.5 g/L glucose, 0.11g/L sodium pyruvate and 2mM glutamine and the cells were grown at 37 ºC in 5 % $CO_2$ atmosphere with 80 % relative humidity.

MDA-MB-231, HT-29, HeLa, HOP-92 and PC-3 cells were plated at 10,000 cells per well (100 μL per well) in 96-well culture dishes and allowed to incubate for 24 h. The old media was removed, cells were washed with PBS (phosphate-buffered saline) and fresh medium was added. A 5 mM solution of inhibitor (in 100% DMSO) was serially diluted in the culture media (8 different concentrations were used) and allowed to incubate for 72 h (MDA-MD-231, HT-29, HOP-92, HL-60, ML2, AML3, PL-21, MOLM-13 and PC-3) or 48h (HeLa and K562). Control cells were treated with the same DMSO concentrations. After the incubation period the medium was removed, cells were washed with PBS to be then incubated with fresh medium containing 86 nM resazurin. Resazurin is reduced to the fluorescent resorufin in the mitochondria: the fluorescence intensity upon incubation in the presence of a living cell culture thus directly correlates with the metabolic viability of the cells. Fluorescence was quantified after 4 hours using a fluorescence microplate reader (Biotek, FLx800™, excitation and emission wavelengths 560 and 590 nm, respectively). The measured fluorescence values were corrected from the control samples containing DMSO.

Leukemia cell lines (K562, HL-60, ML2, AML3, PL-21 and MOLM-13) were seeded at a density of 20,000 cells per well in 100 μL of RMPI media in 96 well microtiter plates. After 24 hours, 12.5 μL of a 10 fold concentrated drug (or DMSO solution for the control) in RMPI media was added in every well. After 48 hour incubation, resazurin was added to every well to obtain a final concentration of 86 μM and, after 3 hours, cell viability was assessed by

measuring the ability of the cells to process resazurin by quantifying the fluorescence using a fluorescence microplate reader (Biotek, FLx800™) as described above.

**Table 9.** Resazurin reduction (percentage of the control) upon incubation of compounds **45**, **48**, **49**, **54**, **58-62** for 48 or 72 h at a concentration of 50 µM in eleven different cancer cell lines. This preliminary screening of toxicity was done as a single experiment. NM indicates no metabolic activity change in comparison to the DMSO-tre ated cells.

| Cmpd | MDA-MD-231 | HT-29 | HeLa | HOP-92 | K562 | HL-60 | ML2 | AML3 | PL-21 | MOLM-13 | PC-3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **45** | NM | NM | NM | 80.3 | 98.3 | 77.1 | 98.3 | 86.8 | 90.2 | 65.2 | 87.7 |
| **48** | 90.7 | NM | 94.8 | 91.0 | NM | 62.8 | NM | 87.3 | 74.7 | 15.1 | 84.3 |
| **49** | NM | NM | 97.6 | 71.9 | NM | 87.8 | 97.3 | 90.0 | 86.0 | NM | 89.1 |
| **54** | 86.7 | NM | 93.2 | 97.8 | 81.3 | 43.2 | 64.1 | 63.0 | 59.4 | 50.2 | 94.1 |
| **58** | NM | NM | 84.5 | 65.3 | 84.8 | 42.1 | 53.2 | 68.9 | 69.2 | 49.4 | 65.0 |
| **59** | 84.3 | NM | 94.2 | 72.8 | NM | 90.3 | NM | 98.7 | NM | 76.3 | 99.9 |
| **60** | 94.4 | NM | 90.4 | 77.3 | NM | 84.7 | NM | 96.9 | 94.3 | 73.2 | 81.8 |
| **61** | 88.2 | NM | 89.3 | 78.3 | 90.8 | 45.7 | 26.5 | 69.9 | 70.5 | 43.0 | 64.1 |
| **62** | NM | NM | NM | NM | 99.2 | 68.6 | NM | 84.2 | 86.3 | 74.8 | 69.5 |

**Table 10.** Metabolic activity values (%) on three different leukemia cell lines at 50 µM compound concentration.[a]

| Cmpd | HL-60 | | ML2 | | MOLM-13 | |
|---|---|---|---|---|---|---|
| | Acid | Ester | Acid | Ester | Acid | Ester |
| **45** | 80 | 78 | 93 | 75 | 71 | 46 (46 µM)[b] |
| **48** | 78 | 89 | 66 | 97 | 19 (33 µM)[b] | 82 |
| **49** | 89 | 89 | 100 | 86 | 89 | 88 |
| **50** | 77 | 85 | 88 | 99 | 73 | 76 |
| **54** | 59 | - | 107 | - | 57 | - |
| **58** | 59 | 96 | 91 | 85 | 60 | 90 |
| **59** | 89 | - | 102 | - | 91 | - |
| **60** | 81 | 74 | 104 | 73 | 85 | 66 |
| **61** | 65 | 73 | 93 | 36 (14.4 µM)[b] | 61 | 11 (5.3 µM)[b] |
| **62** | 82 | 88 | 66 | 60 | 60 | 26 (37 µM)[b] |

[a] Metabolic activity values were determined using resazurin reduction after 72 h incubation with the corresponding compound in comparison to DMSO treated cells at eight different concentrations (100-0.8 µM). The % of metabolic activity at 50 µM compound concentration is shown. The % of metabolic activity is given as the mean of at least three independent experiments. Variability around the mean value was $< 26$ % in all cases. [b] Values in parentheses are $GI_{50}$ values as determined by triplicate experiments with less than 30 % standard error.

### 3.3.12 Comparison table to known CREBBP ligands

**Table 11.** Activity and selectivity of acyl benzene derivatives and comparison with previously reported CREBBP bromodomain ligands.

| Cmpd | LE[a] | LLE | $K_d$ (µM) ITC | $K_d$ (µM) competition binding assay[144-145] | | | | | $\Delta T_m$ (°C)[b] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | CBP | CBP | BRD4(1) | BRD4(2) | BRD2(1,2) | S[c] | CBP | EP300 | BRD4(1) |
| This study | | | | | | | | | | | |
| **45** | 0.35 | 3.7 | 2.0 | 0.77 | >50 | >50 | >50 | >65 | 3.6 | 3.4 | 0.5 |
| **58** | 0.32 | 3.2 | 0.3 | 0.17 | 10 | 36 | 36 | 59 | 4.9 | 5.9 | 1.7 |
| **60** | 0.32 | 4.0 | 0.8 | 0.54 | 26 | >50 | >50 | 48 | 5.0 | 4.6 | 1.3 |
| **61** | 0.28 | 3.8 | – | 0.40 | - | - | - | - | 6.0 | 5.9 | 1.6 |
| Reported by others | | | | | | | | | | | |
| **31**[126] | 0.25 | 2.4 | | 0.20[d] | 0.16[d] | - | - | 0.8 | 7.6 | - | 4.4 |
| **32**[125] | 0.29 | 4.3 | 0.3 | | 1.38[e] | - | - | 4 | 5.4 | - | – |
| **33**[99] | 0.27 | 2.0 | 0.021 | 0.080[d] | 0.85 | 5.2 | - | 40 | 9.7 | 9.7 | 1.8 |

[a] LE = ligand efficiency calculated as (ΔG/number of heavy atoms) is reported in kcal/mol per heavy atom; LLE = lipophilic ligand efficiency (calculated as $pK_d$-*c*logP),[150-151] clogP was calculated using ChemDraw. [b] Median value of the shift in the melting temperature (number of measurements between 9 and 38 per compound/protein pair). The largest SEM was 0.5 °C and most SEM values were below 0.2 °C. [c] Selectivity (S) between the CREBBP and BRD4(1) bromodomains determined by the ratio of $K_d$ values obtained via the competition binding assay. It was not possible to calculate the selectivity using the $K_d$ values obtained via ITC due to the impossibility to reliably fit the titration curves for the BRD4(1) bromodomain. [d] Potency determined by AlphaScreen. [e] Potency determined by Isothermal Titration Calorimetry (ITC).

## 3.4 References


1. Kouzarides, T. Chromatin modifications and their function. *Cell* **2007**, *128*, 693-705.
2. Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561-563.
3. Chung, C. W. Small Molecule Bromodomain Inhibitors: Extending the Druggable Genome. *Progr. Med. Chem.* **2012**, *51*, 1-55.
4. Baylin, S. B.; Jones, P. A. A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer* **2011**, *11*, 726-734.
5. Dawson, M. A.; Kouzarides, T. Cancer Epigenetics: From Mechanism to Therapy. *Cell* **2012**, *150*, 12-27.
6. Pedersen, M. T.; Helin, K. Histone demethylases in development and disease. *Trends Cell Biol.* **2010**, *20*, 662-671.
7. Turner, B. M. Defining an epigenetic code. *Nat. Cell Biol.* **2007**, *9*, 2-6.
8. Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41-45.
9. Jenuwein, T.; Allis, C. D. Translating the histone code. *Science* **2001**, *293*, 1074-1080.
10. Adapted from www.resverlogix.com.

11. Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025-1040.
12. Turner, B. M. Reading signals on the nucleosome with a new nomenclature for modified histones. *Nat. Struct. Mol. Biol.* **2005**, *12*, 110-112.
13. Kouzarides, T. Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* **2000**, *19*, 1176-1179.
14. Conway, S. J. Bromodomains: Are Readers Right for Epigenetic Therapy? *ACS Med. Chem. Lett.* **2012**, *3*, 691-694.
15. Duvic, M.; Vu, J. Vorinostat: a new oral histone deacetylase inhibitor approved for cutaneous T-cell lymphoma. *Expert Opin. Inv. Drug* **2007**, *16*, 1111-1120.
16. Bertino, E. M.; Otterson, G. A. Romidepsin: a novel histone deacetylase inhibitor for cancer. *Expert Opin. Inv. Drug* **2011**, *20*, 1151-1158.
17. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y. C.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E. Selective inhibition of BET bromodomains. *Nature* **2010**, *468*, 1067-1073.
18. Zuber, J.; Shi, J. W.; Wang, E.; Rappaport, A. R.; Herrmann, H.; Sison, E. A.; Magoon, D.; Qi, J.; Blatt, K.; Wunderlich, M.; Taylor, M. J.; Johns, C.; Chicas, A.; Mulloy, J. C.; Kogan, S. C.; Brown, P.; Valent, P.; Bradner, J. E.; Lowe, S. W.; Vakoc, C. R. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **2011**, *478*, 524-U124.
19. Mertz, J. A.; Conery, A. R.; Bryant, B. M.; Sandy, P.; Balasubramanian, S.; Mele, D. A.; Bergeron, L.; Sims, R. J. Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16669-16674.
20. Nicodeme, E.; Jeffrey, K. L.; Schaefer, U.; Beinke, S.; Dewell, S.; Chung, C. W.; Chandwani, R.; Marazzi, I.; Wilson, P.; Coste, H.; White, J.; Kirilovsky, J.; Rice, C. M.; Lora, J. M.; Prinjha, R. K.; Lee, K.; Tarakhovsky, A. Suppression of inflammation by a synthetic histone mimic. *Nature* **2010**, *468*, 1119-1123.
21. Dawson, M. A.; Prinjha, R. K.; Dittmann, A.; Giotopoulos, G.; Bantscheff, M.; Chan, W. I.; Robson, S. C.; Chung, C. W.; Hopf, C.; Savitski, M. M.; Huthmacher, C.; Gudgin, E.; Lugo, D.; Beinke, S.; Chapman, T. D.; Roberts, E. J.; Soden, P. E.; Auger, K. R.; Mirguet, O.; Doehner, K.; Delwel, R.; Burnett, A. K.; Jeffrey, P.; Drewes, G.; Lee, K.; Huntly, B. J. P.; Kouzarides, T. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **2011**, *478*, 529-533.
22. Delmore, J. E.; Issa, G. C.; Lemieux, M. E.; Rahl, P. B.; Shi, J. W.; Jacobs, H. M.; Kastritis, E.; Gilpatrick, T.; Paranal, R. M.; Qi, J.; Chesi, M.; Schinzel, A. C.; McKeown, M. R.; Heffernan, T. P.; Vakoc, C. R.; Bergsagel, P. L.; Ghobrial, I. M.; Richardson, P. G.; Young, R. A.; Hahn, W. C.; Anderson, K. C.; Kung, A. L.; Bradner, J. E.; Mitsiades, C. S. BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc. *Cell* **2011**, *146*, 903-916.
23. Sanchez, R.; Zhou, M. M. The role of human bromodomains in chromatin biology and gene transcription. *Curr. Opin. Drug. Disc.* **2009**, *12*, 659-665.
24. Denis, G. V. Bromodomain Coactivators in Cancer, Obesity, Type 2 Diabetes, and Inflammation. *Discov. Med.* **2010**, *55*, 489-499.

25. Denis, G. V.; Nikolajczyk, B. S.; Schnitzler, G. R. An emerging role for bromodomain-containing proteins in chromatin regulation and transcriptional control of adipogenesis. *FEBS Lett.* **2010**, *584*, 3260-3268.
26. Florence, B.; Faller, D. V. You BET-CHA: A novel family of transcriptional regulators. *Front. Biosci.* **2001**, *6*, D1008-D1018.
27. Muller, S.; Filippakopoulos, P.; Knapp, S. Bromodomains as therapeutic targets. *Expert Rev. Mol. Med.* **2011**, *13*, 1-21.
28. Furdas, S. D.; Carlino, L.; Sippl, W.; Jung, M. Inhibition of bromodomain-mediated protein-protein interactions as a novel therapeutic strategy. *Med. Chem. Commun.* **2012**, *3*, 123-134.
29. Haynes, S. R.; Dollard, C.; Winston, F.; Beck, S.; Trowsdale, J.; Dawid, I. B. The Bromodomain - a Conserved Sequence Found in Human, Drosophila and Yeast Proteins. *Nucleic Acids Res.* **1992**, *20*, 2603-2603.
30. Mujtaba, S.; He, Y.; Zeng, L.; Farooq, A.; Carlson, J. E.; Ott, M.; Verdin, E.; Zhou, M. M. Structural basis of lysine-acetylated HIV-1 tat recognition by PCAF bromodomain. *Mol. Cell* **2002**, *9*, 575-586.
31. Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J. P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Muller, S.; Pawson, T.; Gingras, A. C.; Arrowsmith, C. H.; Knapp, S. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **2012**, *149*, 214-231.
32. Zeng, L.; Zhou, M. M. Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* **2002**, *513*, 124-128.
33. Zhang, Q.; Chakravarty, S.; Ghersi, D.; Zeng, L.; Plotnikov, A. N.; Sanchez, R.; Zhou, M. M. Biochemical Profiling of Histone Binding Selectivity of the Yeast Bromodomain Family. *Plos One* **2010**, *5*.
34. VanDemark, A. P.; Kasten, M. M.; Ferris, E.; Heroux, A.; Hill, C. P.; Cairns, B. R. Autoregulation of the Rsc4 tandem bromodomain by Gcn5 acetylation. *Mol. Cell* **2007**, *27*, 817-828.
35. Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.* **2012**, *11*, 384-400.
36. Vollmuth, F.; Blankenfeldt, W.; Geyer, M. Structures of the Dual Bromodomains of the P-TEFb-activating Protein Brd4 at Atomic Resolution. *J. Biol. Chem.* **2009**, *284*, 36547-36556.
37. Chung, C. W.; Witherington, J. Progress in the Discovery of Small-Molecule Inhibitors of Bromodomain-Histone Interactions. *J. Biomol. Screen.* **2011**, *16*, 1170-1185.
38. Muller, S.; Knapp, S. Discovery of BET bromodomain inhibitors and their role in target validation. *Med. Chem. Commun.* **2014**, *5*, 288-296.
39. Mujtaba, S.; Zeng, L.; Zhou, M. M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **2007**, *26*, 5521-5527.
40. Owen, D. J.; Ornaghi, P.; Yang, J. C.; Lowe, N.; Evans, P. R.; Ballario, P.; Neuhaus, D.; Filetici, P.; Travers, A. A. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.* **2000**, *19*, 6141-6149.
41. Filippakopoulos, P.; Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug Discov.* **2014**, *13*, 339-358.

42. Hewings, D. S.; Rooney, T. P. C.; Jennings, L. E.; Hay, D. A.; Schofield, C. J.; Brennan, P. E.; Knapp, S.; Conway, S. J. Progress in the Development and Application of Small Molecule Inhibitors of Bromodomain-Acetyl-lysine Interactions. *J. Med. Chem.* **2012**, *55*, 9393-9413.
43. Steiner, S.; Magno, A.; Huang, D. Z.; Caflisch, A. Does bromodomain flexibility influence histone recognition? *FEBS Lett.* **2013**, *587*, 2158-2163.
44. Dhalluin, C.; Carlson, J. E.; Zeng, L.; He, C.; Aggarwal, A. K.; Zhou, M. M. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **1999**, *399*, 491-496.
45. Chung, C. W.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A.; Boursier, E. V.; Flynn, H.; Bouillot, A. M.; Bamborough, P.; Brusq, J. M. G.; Gellibert, F. J.; Jones, E. J.; Riou, A. M.; Homes, P.; Martin, S. L.; Uings, I. J.; Toum, J.; Clement, C. A.; Boullay, A. B.; Grimley, R. L.; Blande, F. M.; Prinjha, R. K.; Lee, K.; Kirilovsky, J.; Nicodeme, E. Discovery and Characterizatlion of Small Molecule Inhibitors of the BET Family Bromodomains. *J. Med. Chem.* **2011**, *54*, 3827-3838.
46. Zeng, L.; Li, J. M.; Muller, M.; Yan, S.; Mujtaba, S.; Pan, C. F.; Wang, Z. Y.; Zhou, M. M. Selective small molecules blocking HIV-1 Tat and coactivator PCAF association. *J. Am. Chem. Soc.* **2005**, *127*, 2376-2377.
47. Jennings, L. E.; Measures, A. R.; Wilson, B. G.; Conway, S. J. Progress in the development and application of small molecule inhibitors of bromodomain-acetyl-lysine interactions. *Future Med. Chem.* **2014**, *6*, 179-204.
48. Zhou, M. S.; Huang, K.; Jung, K. J.; Cho, W. K.; Klase, Z.; Kashanchi, F.; Pise-Masison, C. A.; Brady, J. N. Bromodomain Protein Brd4 Regulates Human Immunodeficiency Virus Transcription through Phosphorylation of CDK9 at Threonine 29. *J. Virol.* **2009**, *83*, 1036-1044.
49. Lin, A.; Wang, S.; Nguyen, T.; Shire, K.; Frappier, L. The EBNA1 Protein of Epstein-Barr Virus Functionally Interacts with Brd4. *J. Virol.* **2008**, *82*, 12009-12019.
50. You, J. X.; Srinivasan, V.; Denis, G. V.; Harrington, W. J.; Ballestas, M. E.; Kaye, K. M.; Howley, P. M. Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen interacts with bromodomain protein Brd4 on host mitotic chromosomes. *J. Virol.* **2006**, *80*, 8909-8919.
51. Abbate, E. A.; Voitenleitner, C.; Botchan, M. R. Structure of the papillomavirus DNA-tethering complex E2 : Brd4 and a peptide that ablates HPV chromosomal association. *Mol. Cell* **2006**, *24*, 877-889.
52. Huang, B.; Yang, X. D.; Zhou, M. M.; Ozato, K.; Chen, L. F. Brd4 Coactivates Transcriptional Activation of NF-kappa B via Specific Binding to Acetylated RelA. *Mol. Cell. Biol.* **2009**, *29*, 1375-1387.
53. Pasparakis, M. Regulation of tissue homeostasis by NF-kappa B signalling: implications for inflammatory diseases. *Nat. Rev. Immunol.* **2009**, *9*, 778-788.
54. Yang, Z. Y.; He, N. H.; Zhou, Q. Brd4 recruits P-TER to chromosomes at late mitosis to promote G(1) gene expression and cell cycle progression. *Mol. Cell. Biol.* **2008**, *28*, 967-976.
55. Dey, A.; Nishiyama, A.; Karpova, T.; McNally, J.; Ozato, K. Brd4 Marks Select Genes on Mitotic Chromatin and Directs Postmitotic Transcription. *Mol. Biol. Cell* **2009**, *20*, 4899-4909.

56. Yang, X. J. Multisite protein modi. cation and intramolecular signaling. *Oncogene* **2005**, *24*, 1653-1662.
57. Yang, Z. Y.; Yik, J. H. N.; Chen, R. C.; He, N. H.; Jang, M. K.; Ozato, K.; Zhou, Q. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein brd4. *Mol. Cell* **2005**, *19*, 535-545.
58. Phelps, M. A.; Lin, T. S.; Johnson, A. J.; Hurh, E.; Rozewski, D. M.; Farley, K. L.; Wu, D.; Blum, K. A.; Fischer, B.; Mitchell, S. M.; Moran, M. E.; Brooker-McEldowney, M.; Heerema, N. A.; Jarjoura, D.; Schaaf, L. J.; Byrd, J. C.; Grever, M. R.; Dalton, J. T. Clinical response and pharmacokinetics from a phase 1 study of an active dosing schedule of flavopiridol in relapsed chronic lymphocytic leukemia. *Blood* **2009**, *113*, 2637-2645.
59. French, C. A.; Miyoshi, I.; Kubonishi, I.; Grier, H. E.; Perez-Atayde, A. R.; Fletcher, J. A. BRD4-NUT fusion oncogene: A novel mechanism in aggressive carcinoma. *Cancer Res.* **2003**, *63*, 304-307.
60. Crawford, N. P. S.; Alsarraj, J.; Lukes, L.; Walker, R. C.; Officewala, J. S.; Yang, H. H.; Lee, M. P.; Ozato, K.; Hunter, K. W. Bromodomain 4 activation predicts breast cancer survival. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6380-6385.
61. Zhao, Y. J.; Yang, C. Y.; Wang, S. M. The Making of I-BET762, a BET Bromodomain Inhibitor Now in Clinical Development. *J. Med. Chem.* **2013**, *56*, 7498-7500.
62. www.clinicaltrials.gov, accessed on 18/10/2014.
63. Filippakopoulos, P.; Picaud, S.; Fedorov, O.; Keller, M.; Wrobel, M.; Morgenstern, O.; Bracher, F.; Knapp, S. Benzodiazepines and benzotriazepines as protein interaction inhibitors targeting bromodomains of the BET family. *Biorg. Med. Chem.* **2012**, *20*, 1878-1886.
64. Zhang, G. T.; Liu, R. J.; Zhong, Y. F.; Plotnikov, A. N.; Zhang, W. J.; Zeng, L.; Rusinova, E.; Gerona-Nevarro, G.; Moshkina, N.; Joshua, J.; Chuang, P. Y.; Ohlmeyer, M.; He, J. C.; Zhou, M. M. Down-regulation of NF-kappa B transcriptional activity in HIV-associated kidney disease by BRD4 inhibition (vol 287, pg 28840, 2012). *J. Biol. Chem.* **2012**, *287*, 38956-38956.
65. Gehling, V. S.; Hewitt, M. C.; Vaswani, R. G.; Leblanc, Y.; Cote, A.; Nasveschuk, C. G.; Taylor, A. M.; Harmange, J. C.; Audia, J. E.; Pardo, E.; Joshi, S.; Sandy, P.; Mertz, J. A.; Sims, R. J.; Bergeron, L.; Bryant, B. M.; Bellon, S.; Poy, F.; Jayaram, H.; Sankaranarayanan, R.; Yellapantula, S.; Srinivasamurthy, N. B.; Birudukota, S.; Albrecht, B. K. Discovery, Design, and Optimization of Isoxazole Azepine BET Inhibitors. *ACS Med. Chem. Lett.* **2013**, *4*, 835-840.
66. Albretcht, B. K.; Audia, J. E.; Cote, A.; Gehling, V. S.; Harmange, J.-C.; Hewitt, M. C.; Leblanc, Y.; Naveschuk, C. G.; Taylor, A. M.; Vaswani, R. G. Bromodomain inhibitors and uses thereof. 2012.
67. Albrecht, B. K.; Gehling, V. S.; Harmange, J. C.; Hewitt, M. C.; Taylor, A. M. Bromodomain inhibitors and uses thereof. 2012.
68. Hewings, D. S.; Wang, M. H.; Philpott, M.; Fedorov, O.; Uttarkar, S.; Filippakopoulos, P.; Picaud, S.; Vuppusetty, C.; Marsden, B.; Knapp, S.; Conway, S. J.; Heightman, T. D. 3,5-Dimethylisoxazoles Act As Acetyl-lysine-mimetic Bromodomain Ligands. *J. Med. Chem.* **2011**, *54*, 6761-6770.
69. Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, M. D.; Chung, C. W. Fragment-Based Discovery of Bromodomain Inhibitors Part 2: Optimization of Phenylisoxazole Sulfonamides. *J. Med. Chem.* **2012**, *55*, 587-596.

70. Mirguet, O.; Lamotte, Y.; Donche, F.; Toum, J.; Gellibert, F.; Bouillot, A.; Gosmini, R.; Nguyen, V. L.; Delannee, D.; Seal, J.; Blandel, F.; Boullay, A. B.; Boursier, E.; Martin, S.; Brusq, J. M.; Krysa, G.; Riou, A.; Tellier, R.; Costaz, A.; Huet, P.; Dudit, Y.; Trottet, L.; Kirilovsky, J.; Nicodeme, E. From ApoA1 upregulation to BET family bromodomain inhibition: Discovery of I-BET151. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2963-2967.
71. Mirguet, O.; Gosmini, R.; Toum, J.; Clement, C. A.; Barnathan, M.; Brusq, J. M.; Mordaunt, J. E.; Grimes, R. M.; Crowe, M.; Pineau, O.; Ajakane, M.; Daugan, A.; Jeffrey, P.; Cutler, L.; Haynes, A. C.; Smithers, N. N.; Chung, C. W.; Bamborough, P.; Uings, I. J.; Lewis, A.; Witherington, J.; Parr, N.; Prinjha, R. K.; Nicodeme, E. Discovery of Epigenetic Regulator I-BET762: Lead Optimization to Afford a Clinical Candidate Inhibitor of the BET Bromodomains. *J. Med. Chem.* **2013**, *56*, 7501-7515.
72. Seal, J.; Lamotte, Y.; Donche, F.; Bouillot, A.; Mirguet, O.; Gellibert, F.; Nicodeme, E.; Krysa, G.; Kirilovsky, J.; Beinke, S.; McCleary, S.; Rioja, I.; Bamborough, P.; Chung, C. W.; Gordon, L.; Lewis, T.; Walker, A. L.; Cutler, L.; Lugo, D.; Wilson, D. M.; Witherington, J.; Lee, K.; Prinjha, R. K. Identification of a novel series of BET family bromodomain inhibitors: Binding mode and profile of I-BET151 (GSK1210151A). *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2968-2972.
73. Hay, D.; Fedorov, O.; Filippakopoulos, P.; Martin, S.; Philpott, M.; Picaud, S.; Hewings, D. S.; Uttakar, S.; Heightman, T. D.; Conway, S. J.; Knapp, S.; Brennan, P. E. The design and synthesis of 5-and 6-isoxazolylbenzimidazoles as selective inhibitors of the BET bromodomains. *Med. Chem. Commun.* **2013**, *4*, 140-144.
74. Picaud, S.; Da Costa, D.; Thanasopoulou, A.; Filippakopoulos, P.; Fish, P. V.; Philpott, M.; Fedorov, O.; Brennan, P.; Bunnage, M. E.; Owen, D. R.; Bradner, J. E.; Taniere, P.; O'Sullivan, B.; Muller, S.; Schwaller, J.; Stankovic, T.; Knapp, S. PFI-1, a Highly Selective Protein Interaction Inhibitor, Targeting BET Bromodomains. *Cancer Res.* **2013**, *73*, 3336-3346.
75. Fish, P. V.; Filippakopoulos, P.; Bish, G.; Brennan, P. E.; Bunnage, M. E.; Cook, A. S.; Federov, O.; Gerstenberger, B. S.; Jones, H.; Knapp, S.; Marsden, B.; Nocka, K.; Owen, D. R.; Philpott, M.; Picaud, S.; Primiano, M. J.; Ralph, M. J.; Sciammetta, N.; Trzupek, J. D. Identification of a Chemical Probe for Bromo and Extra C-Terminal Bromodomain Inhibition through Optimization of a Fragment-Derived Hit. *J. Med. Chem.* **2012**, *55*, 9831-9837.
76. DeMont, E. H.; Gosmini, R. L. Tetrahydroquinolines Derivatives as Bromodomain Inhibitors. 2011.
77. Chung, C. W.; Dean, A. W.; Woolven, J. M.; Bamborough, P. Fragment-Based Discovery of Bromodomain Inhibitors Part 1: Inhibitor Binding Modes and Implications for Lead Discovery. *J. Med. Chem.* **2012**, *55*, 576-586.
78. Albrecht, B. K.; Harmange, J. C.; Cote, A.; Taylor, A. M. Bromodomain inhibitors and uses thereof. 2012.
79. Wang, L.; Pratt, J. K.; Mcdaniel, K. F. Bromodomain inhibitors. 2013.
80. Wang, L.; Pratt, J. K.; Mcdaniel, K. F.; Dai, Y.; Fidanze, S. D.; Hasvold, L.; Holms, J. H.; Kati, W. M.; Liu, D.; Mantei, R. A.; McClellan, W. J.; Sheppard, G. S.; Wada, C. K. Bromdomain inhibitors. 2013.
81. Ito, T.; Umehara, T.; Sasaki, K.; Nakamura, Y.; Nishino, N.; Terada, T.; Shirouzu, M.; Padmanabhan, B.; Yokoyama, S.; Ito, A.; Yoshida, M. Real-Time Imaging of Histone H4K12-

Specific Acetylation Determines the Modes of Action of Histone Deacetylase and Bromodomain Inhibitors. *Chem. Biol.* **2011**, *18*, 495-507.

82. Zhao, L. L.; Cao, D. Y.; Chen, T. T.; Wang, Y. Q.; Miao, Z. H.; Xu, Y. C.; Chen, W. Y.; Wang, X.; Li, Y. D.; Du, Z. Y.; Xiong, B.; Li, J.; Xu, C. Y.; Zhang, N. X.; He, J. H.; Shen, J. K. Fragment-Based Drug Discovery of 2-Thiazolidinones as Inhibitors of the Histone Reader BRD4 Bromodomain. *J. Med. Chem.* **2013**, *56*, 3833-3851.

83. Miyoshi, S.; Ooike, S.; Iwata, K.; Hikawa, H.; Sugahara, K. Antitumor agent. 2009.

84. Barter, P. J.; Nicholls, S.; Rye, K. A.; Anantharamaiah, G. M.; Navab, M.; Fogelman, A. M. Antiinflammatory properties of HDL. *Circul. Res.* **2004**, *95*, 764-772.

85. Bradner, J.; Brown, J.; Plutzky, J. Compositions and methods for modulating metabolism. 2011.

86. Bradner, J. E.; Shi, J.; Vakoc, C. R.; Zuber, J. Compositions and methods for treating leukemia. 2011.

87. Shimamura, T.; Chen, Z.; Soucheray, M.; Carretero, J.; Kikuchi, E.; Tchaicha, J. H.; Gao, Y. D.; Cheng, K. A.; Cohoon, T. J.; Qi, J.; Akbay, E.; Kimmelman, A. C.; Kung, A. L.; Bradner, J. E.; Wong, K. K. Efficacy of BET Bromodomain Inhibition in Kras-Mutant Non-Small Cell Lung Cancer. *Clin. Cancer. Res.* **2013**, *19*, 6183-6192.

88. NIH http://clinicaltrials.gov: identifier NCT01587703.

89. Hewings, D. S.; Fedorov, O.; Filippakopoulos, P.; Martin, S.; Picaud, S.; Tumber, A.; Wells, C.; Olcina, M. M.; Freeman, K.; Gill, A.; Ritchie, A. J.; Sheppard, D. W.; Russell, A. J.; Hammond, E. M.; Knapp, S.; Brennan, P. E.; Conway, S. J. Optimization of 3,5-Dimethylisoxazole Derivatives as Potent Bromodomain Ligands. *J. Med. Chem.* **2013**, *56*, 3217-3227.

90. Gallenkamp, D.; Gelato, K. A.; Haendler, B.; Weinmann, H. Bromodomains and Their Pharmacological Inhibitors. *ChemMedChem* **2014**, *9*, 438-464.

91. Fish, P. V.; Cook, A. S.; Phillips, C.; Bent, A. F.; Mills, J. e. J.; Sciammetta, N. Novel heterocyclic compounds as bromodomain inhibitors. 2013.

92. Wyce, A.; Ganji, G.; Smitheman, K. N.; Chung, C. W.; Korenchuk, S.; Bai, Y. C.; Barbash, O.; Le, B. C.; Craggs, P. D.; McCabe, M. T.; Kennedy-Wilson, K. M.; Sanchez, L. V.; Gosmini, R. L.; Parr, N.; McHugh, C. F.; Dhanak, D.; Prinjha, R. K.; Auger, K. R.; Tummino, P. J. BET Inhibition Silences Expression of MYCN and BCL2 and Induces Cytotoxicity in Neuroblastoma Tumor Models. *Plos One* **2013**, *8*.

93. Vidler, L. R.; Filippakopoulos, P.; Fedorov, O.; Picaud, S.; Martin, S.; Tomsett, M.; Woodward, H.; Brown, N.; Knapp, S.; Hoelder, S. Discovery of Novel Small-Molecule Inhibitors of BRD4 Using Structure-Based Virtual Screening. *J. Med. Chem.* **2013**, *56*, 8073-8088.

94. Lucas, X.; Wohlwend, D.; Hugle, M.; Schmidtkunz, K.; Gerhardt, S.; Schule, R.; Jung, M.; Einsle, O.; Gunther, S. 4-Acyl Pyrroles: Mimicking Acetylated Lysines in Histone Code Reading. *Angew. Chem. Int. Ed.* **2013**, *52*, 14055-14059.

95. Zhao, H. T.; Gartenmann, L.; Dong, J.; Spiliotopoulos, D.; Caflisch, A. Discovery of BRD4 bromodomain inhibitors by fragment-based high-throughput docking. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2493-2496.

96. Kolb, P.; Kipouros, C. B.; Huang, D. Z.; Caflisch, A. Structure-based tailoring of compound libraries for high-throughput screening: Discovery of novel EphB4 kinase inhibitors. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 11-18.

97. Ciceri, P.; Muller, S.; O'Mahony, A.; Fedorov, O.; Filippakopoulos, P.; Hunt, J. P.; Lasater, E. A.; Pallares, G.; Picaud, S.; Wells, C.; Martin, S.; Wodicka, L. M.; Shah, N. P.; Treiber, D. K.; Knapp, S. Dual kinase-bromodomain inhibitors for rationally designed polypharmacology. *Nat. Chem. Biol.* **2014**, *10*, 305-312.
98. Ember, S. W. J.; Zhu, J. Y.; Olesen, S. H.; Martin, M. P.; Becker, A.; Berndt, N.; Georg, G. I.; Schonbrunn, E. Acetyl-lysine Binding Site of Bromodomain-Containing Protein 4 (BRD4) Interacts with Diverse Kinase Inhibitors. *ACS Chem. Biol.* **2014**, *9*, 1160-1171.
99. Hay, D. A.; Fedorov, O.; Martin, S.; Singleton, D. C.; Tallant, C.; Wells, C.; Picaud, S.; Philpott, M.; Monteiro, O. P.; Rogers, C. M.; Conway, S. J.; Rooney, T. P. C.; Tumber, A.; Yapp, C.; Filippakopoulos, P.; Bunnage, M. E.; Muller, S.; Knapp, S.; Schofield, C. J.; Brennan, P. E. Discovery and Optimization of Small-Molecule Ligands for the CBP/p300 Bromodomains. *J. Am. Chem. Soc.* **2014**, *136*, 9308-9319.
100. Bannister, A. J.; Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **1996**, *384*, 641-643.
101. Ogryzko, V. V.; Schiltz, R. L.; Russanova, V.; Howard, B. H.; Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **1996**, *87*, 953-959.
102. Wang, F.; Marshall, C. B.; Ikura, M. Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. *Cell. Mol. Life Sci.* **2013**, *70*, 3989-4008.
103. Bedford, D. C.; Kasper, L. H.; Fukuyama, T.; Brindle, P. K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics-Us* **2010**, *5*, 9-15.
104. Ait-Si-Ali, S.; Polesskaya, A.; Filleur, S.; Ferreira, R.; Duquet, A.; Robin, P.; Vervish, A.; Trouche, D.; Cabon, F.; Harel-Bellan, A. CBP/p300 histone acetyl-transferase activity is important for the G1/S transition. *Oncogene* **2000**, *19*, 2430-2437.
105. Iyer, N. G.; Xian, J.; Chin, S. F.; Bannister, A. J.; Daigo, Y.; Aparicio, S.; Kouzarides, T.; Caldas, C. p300 is required for orderly G1/S transition in human cancer cells. *Oncogene* **2007**, *26*, 21-29.
106. Mujtaba, S.; He, Y.; Zeng, L.; Yan, S.; Plotnikova, O.; Sachchidanand; Sanchez, R.; Zeleznik-Le, N. J.; Ronai, Z.; Zhou, M. M. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol. Cell* **2004**, *13*, 251-263.
107. Prives, C.; Hall, P. A. The P53 pathway. *J. Pathol.* **1999**, *187*, 112-126.
108. Coutts, A. S.; La Thangue, N. B. The p53 response: Emerging levels of co-factor complexity. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 778-785.
109. Alarcon-Vargas, D.; Ronai, Z. p53-Mdm2 - the affair that never ends. *Carcinogenesis* **2002**, *23*, 541-547.
110. Vogelstein, B.; Lane, D.; Levine, A. J. Surfing the p53 network. *Nature* **2000**, *408*, 307-310.
111. Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **1997**, *88*, 323-331.
112. Nayak, S. K.; Panesar, P. S.; Kumar, H. p53-Induced Apoptosis and Inhibitors of p53. *Curr. Med. Chem.* **2009**, *16*, 2627-2640.
113. Li, J. D.; Ghiani, C. A.; Kim, J. Y.; Liu, A. X.; Sandoval, J.; DeVellis, J.; Casaccia-Bonnefil, P. Inhibition of p53 transcriptional activity: A potential target for future

development of therapeutic strategies for primary demyelination. *J. Neurosci.* **2008**, *28*, 6118-6127.
114. Duan, W. Z.; Zhu, X. X.; Ladenheim, B.; Yu, Q. S.; Guo, Z. H.; Oyler, J.; Cutler, R. G.; Cadet, J. L.; Greig, N. H.; Mattson, M. P. p53 inhibitors preserve dopamine neurons and motor function in experimental parkinsonism. *Ann. Neurol.* **2002**, *52*, 597-606.
115. Hong, L. Z.; Zhao, X. Y.; Zhang, H. L. p53-mediated neuronal cell death in ischemic brain injury. *Neurosci. Bull.* **2010**, *26*, 232-240.
116. Taishiro, J.; Kikuchi, S.; Shinpo, K.; Kishirnoto, R.; Tsuji, S.; Sasaki, H. Role of p53 in neurotoxicity induced by the endoplasmic reticulum stress agent tunicamycin in cultures of rat organotypic slice spinal cord. *J. Neurosci. Res.* **2007**, *85*, 395-401.
117. Pietrancosta, N.; Garino, C.; Laras, Y.; Quelever, G.; Pierre, P.; Clavarino, G.; Kraus, J. L. Are p53 inhibitors potentially useful therapeutics? *Drug. Develop. Res.* **2005**, *65*, 43-49.
118. Bedford, D. C.; Brindle, P. K. Is histone acetylation the most important physiological function for CBP and p300? *Aging-Us* **2012**, *4*, 247-255.
119. Lavau, C.; Du, C. C.; Thirman, M.; Zeleznik-Le, N. Chromatin-related properties of CBP fused to MLL generate a myelodysplastic-like syndrome that evolves into myeloid leukemia. *EMBO J.* **2000**, *19*, 4655-4664.
120. Pasqualucci, L.; Trifonov, V.; Fabbri, G.; Ma, J.; Rossi, D.; Chiarenza, A.; Wells, V. A.; Grunn, A.; Messina, M.; Elliot, O.; Chan, J.; Bhagat, G.; Chadburn, A.; Gaidano, G.; Mullighan, C. G.; Rabadan, R.; Dalla-Favera, R. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* **2011**, *43*, 830-U833.
121. Dekker, F. J.; Haisma, H. J. Histone acetyl transferases as emerging drug targets. *Drug Discov. Today* **2009**, *14*, 942-948.
122. Heery, D. M.; Fischer, P. M. Pharmacological targeting of lysine acetyltransferases in human disease: a progress report. *Drug Discov. Today* **2007**, *12*, 88-99.
123. Sachchidanand; Resnick-Silverman, L.; Yan, S.; Mutjaba, S.; Liu, W. J.; Zeng, L.; Manfredi, J. J.; Zhou, M. M. Target structure-based discovery of small molecules that block human p53 and CREB binding protein association. *Chem. Biol.* **2006**, *13*, 81-90.
124. Borah, J. C.; Mujtaba, S.; Karakikes, I.; Zeng, L.; Muller, M.; Patel, J.; Moshkina, N.; Morohashi, K.; Zhang, W. J.; Gerona-Navarro, G.; Hajjar, R. J.; Zhou, M. M. A Small Molecule Binding to the Coactivator CREB-Binding Protein Blocks Apoptosis in Cardiomyocytes. *Chem. Biol.* **2011**, *18*, 531-541.
125. Rooney, T. P. C.; Filippakopoulos, P.; Fedorov, O.; Picaud, S.; Cortopassi, W. A.; Hay, D. A.; Martin, S.; Tumber, A.; Rogers, C. M.; Philpott, M.; Wang, M. H.; Thompson, A. L.; Heightman, T. D.; Pryde, D. C.; Cook, A.; Paton, R. S.; Muller, S.; Knapp, S.; Brennan, P. E.; Conway, S. J. A Series of Potent CREBBP Bromodomain Ligands Reveals an Induced-Fit Pocket Stabilized by a Cation-pi Interaction. *Angew. Chem. Int. Ed.* **2014**, *53*, 6126-6130.
126. Fedorov, O.; Lingard, H.; Wells, C.; Monteiro, O. P.; Picaud, S.; Keates, T.; Yapp, C.; Philpott, M.; Martin, S. J.; Felletar, I.; Marsden, B. D.; Filippakopoulos, P.; Muller, S.; Knapp, S.; Brennan, P. E. [1,2,4]Triazolo[4,3-a]phthalazines: Inhibitors of Diverse Bromodomains. *J. Med. Chem.* **2014**, *57*, 462-476.
127. SGC-CBP30: a CREBBP/EP300-selective chemical probe: http://www.thesgc.org/chemical-probes/SGC-CBP30.
128. ICPB112: a CREBBP/EP300-selective chemical probe: http://www.thesgc.org/chemical-probes/ICBP112.

129. Vidler, L. R.; Brown, N.; Knapp, S.; Hoelder, S. Druggability Analysis and Structural Classification of Bromodomain Acetyl-lysine Binding Sites. *J. Med. Chem.* **2012**, *55*, 7346-7359.
130. Bromosporine: http://www.thesgc.org/chemical-probes/bromosporine.
131. Ferguson, F. M.; Fedorov, O.; Chaikuad, A.; Philpott, M.; Muniz, J. R. C.; Felletar, I.; von Delft, F.; Heightman, T.; Knapp, S.; Abell, C.; Ciulli, A. Targeting Low-Druggability Bromodomains: Fragment Based Screening and Inhibitor Design against the BAZ2B Bromodomain. *J. Med. Chem.* **2013**, *56*, 10183-10187.
132. GSK2801: A selective chemical probe for BAZ2B/A bromodomains: http://www.thesgc.org/chemical-probes/GSK2801.
133. PFI-3: Selective chemical probe for SMARCA bromodomains: http://www.thesgc.org/chemical-probes/PFI-3.
134. Demont, E. H.; Bamborough, P.; Chung, C.; Craggs, P. D.; Fallon, D.; Gordon, L. J.; Grandi, P.; Hobbs, C. I.; Hussain, J.; Jones, E. J.; Le Gall, A.; Michon, A.; Mitchell, D. J.; Prinjha, R. K.; Roberts, A. D.; Sheppard, R. J.; Watson, R. J. 1,3-Dimethyl Benzimidazolones Are Potent, Selective Inhibitors of the BRPF1 Bromodomain. *ACS Med. Chem. Lett.* **2014**, *dx.doi.org/10.1021/ml5002932*.
135. Spiliotopoulos, D.; Caflisch, A. Molecular Dynamics Simulations of Bromodomains Reveal Binding-Site Flexibility and Multiple Binding Modes of the Natural Ligand Acetyl-Lysine. *Isr. J. Chem.* **2014**, *54*, 1084-1092.
136. Kolb, P.; Kipouros, C. B.; Huang, D.; Caflisch, A. Structure-based tailoring of compound libraries for high-throughput screening: Discovery of novel EphB4 kinase inhibitors. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 11-18.
137. Lafleur, K.; Huang, D. Z.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-Based Optimization of Potent and Selective Inhibitors of the Tyrosine Kinase Erythropoietin Producing Human Hepatocellular Carcinoma Receptor B4 (EphB4). *J. Med. Chem.* **2009**, *52*, 6433-6446.
138. Zhao, H.; Dong, J.; Lafleur, K.; Nevado, C.; Caflisch, A. Discovery of a Novel Chemotype of Tyrosine Kinase Inhibitors by Fragment-Based Docking and Molecular Dynamics. *ACS Med. Chem. Lett.* **2012**, *3*, 834-838.
139. Kolb, P.; Caflisch, A. Automatic and Efficient Decomposition of Two-Dimensional Structures of Small Molecules for Fragment-Based High-Throughput Docking. *J. Med. Chem.* **2006**, *49*, 7384-7392.
140. Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhardt, C.; Caflisch, A. Exhaustive Docking of Molecular Fragments With Electrostatic Solvation. *Proteins* **1999**, *37*, 88-105.
141. Majeux, N.; Scarsi, M.; Caflisch, A. Efficient electrostatic solvation model for protein-fragment docking. *Proteins: Struct., Funct., Bioinf.* **2001**, *42*, 256-268.
142. Scarsi, M.; Apostolakis, J.; Caflisch, A. Continuum Electrostatic Energies of Macromolecules in Aqueous Solutions. *J. Phys. Chem. A* **1997**, *101*, 8098-8106.
143. Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455-461.
144. Fabian, M. A.; Biggs, W. H.; Treiber, D. K.; Atteridge, C. E.; Azimioara, M. D.; Benedetti, M. G.; Carter, T. A.; Ciceri, P.; Edeen, P. T.; Floyd, M.; Ford, J. M.; Galvin, M.; Gerlach, J. L.; Grotzfeld, R. M.; Herrgard, S.; Insko, D. E.; Insko, M. A.; Lai, A. G.; Lelias, J.

M.; Mehta, S. A.; Milanov, Z. V.; Velasco, A. M.; Wodicka, L. M.; Patel, H. K.; Zarrinkar, P. P.; Lockhart, D. J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **2005**, *23*, 329-336.

145. Quinn, E.; Wodicka, L.; Ciceri, P.; Pallares, G.; Pickle, E.; Torrey, A.; Floyd, M.; Hunt, J.; Treiber, D. Abstract 4238: BROMOscan - a high throughput, quantitative ligand binding platform identifies best-in-class bromodomain inhibitors from a screen of mature compounds targeting other protein classes. *Cancer Res.* **2013**, *73*, 4238.

146. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545-1614.

147. Im, W.; Beglov, D.; Roux, B. Continuum Solvation Model: computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation. *Comput. Phys. Commun.* **1998**, *111*, 59-75.

148. The finite-difference Poisson calculation was repeated on the energy minimized crystal structure. Again, the electrostatic interaction between the carboxyl group of compound **45** and the Arg1173 side chain guanidinium (-12.4 kcal/mol) is about half of the total electrostatic interaction energy (-25.0 kcal/mol ).

149. The weak signal obtained for BRD4(1) does not allow a reliable fit of the titration curve (See Figure 31 in the experimental section).

150. Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **2004**, *9*, 430-431.

151. Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* **2007**, *6*, 881-890.

152. Borrow, J.; Stanton, V. P.; Andresen, J. M.; Becher, R.; Behm, F. G.; Chaganti, R. S. K.; Civin, C. I.; Disteche, C.; Dube, I.; Frischauf, A. M.; Horsman, D.; Mitelman, F.; Volinia, S.; Watmore, A. E.; Housman, D. E. The translocation t(8;l6)(p11, p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB binding protein. *Nat. Genet.* **1996**, *14*, 33-41.

153. Sobulo, O. M.; Borrow, J.; Tomek, R.; Reshmi, S.; Harden, A.; Schlegelberger, B.; Housman, D.; Doggett, N. A.; Rowley, J. D.; ZeleznikLe, N. J. MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8732-8737.

154. Goodman, R. H.; Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **2000**, *14*, 1553-1577.

155. The discrepancy observed for compound **64** might stem from the different chemical entities used in the calculation (head fragment only) and experiment (whole compound), and the approximations intrinsic to rigid protein docking.

156. Chung, C.-W.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A.; Boursier, E. V.; Flynn, H.; Bouillot, A. M.; Bamborough, P.; Brusq, J.-M. G.; Gellibert, F. J.; Jones, E. J.; Riou, A. M.; Homes, P.; Martin, S. L.; Uings, I. J.; Toum, J.; Clément, C. A.; Boullay, A.-B.; Grimley, R. L.; Blandel, F. M.; Prinjha, R. K.; Lee, K.; Kirilovsky, J.; Nicodeme, E. Discovery and Characterization of

Small Molecule Inhibitors of the BET Family Bromodomains. *J. Med. Chem.* **2011**, *54*, 3827-3838.

157. Scarsi, M.; Apostolakis, J.; Caflisch, A. Comparison of a GB Solvation Model with Explicit Solvent Simulations: Potentials of Mean Force and Conformational Preferences of Alanine Dipeptide and 1,2-Dichloroethane. *J. Phys. Chem. B* **1998**, *102*, 3637-3641.

158. Zhao, H.; Huang, D. Hydrogen Bonding Penalty upon Ligand Binding. *PLoS ONE* **2011**, *6*, e19923.

159. Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L. MATCH: An Atom-Typing Toolset for Molecular Mechanics Force Fields. *J. Comput. Chem.* **2012**, *33*, 189-202.

160. Huang, J.; MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135-2145.

161. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A. D. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671-690.

162. Scarsi, M.; Majeux, N.; Caflisch, A. Hydrophobicity at the surface of proteins. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 565-575.

163. Baldwin, J. E.; Cianciosi, S. J. Syntheses of Racemic and Both Chiral Forms of Cyclopropane-1,2-D(2) and Cyclopropane-1-C-13-1,2,3-D(3). *J. Am. Chem. Soc.* **1992**, *114*, 9401-9408.

164. Benatti, L.; Fariello, R.; Salvati, P.; Pellicciari, R.; Caccia, C. *Newron Pharm. Spa.* **2004**, *EP1424333 (A1)*.

165. Voitekhovich, S. V.; Vorobiov, A. N.; Gaponik, P. N.; Ivashkevich, O. A. Synthesis of new functionally substituted 1-R-tetrazoles and their 5-amino derivatives. *Khim Geterotsikl+* **2005**, 1174-1179.

166. Bridger, G.; Skerlj, R.; Kaller, A.; Harwig, C.; Bogucki, D.; Wilson, T. R.; Crawford, J.; McEachern, E. J.; Atsma, B.; Nan, S.; Zhou, Y.; Schols, S.; Smith, C. D.; Fluri, M. R. D. Chemokine receptor binding heterocyclic compounds **2002**, *US2002147192 (A1)*.

167. Wydysh, E. A.; Medghalchi, S. M.; Vadlamudi, A.; Townsend, C. A. Design and Synthesis of Small Molecule Glycerol 3-Phosphate Acyltransferase Inhibitors. *J. Med. Chem.* **2009**, *52*, 3317-3327.

168. Feutrill, J.; Leriche, C.; Middlemiss, D. *Fovea Pharmaceuticals* **2013**, *WO2013037705 (A2)*.

169. Meng, L.; Cheng, Q. G.; Kim, C.; Gao, W. Y.; Wojtas, L.; Chen, Y. S.; Zaworotko, M. J.; Zhang, X. P.; Ma, S. Q. Crystal Engineering of a Microporous, Catalytically Active fcu Topology MOF Using a Custom-Designed Metalloporphyrin Linker. *Angew. Chem. Int. Ed.* **2012**, *51*, 10082-10085.

170. Zhao, G. Z.; Chen, L. J.; Wang, W.; Zhang, J.; Yang, G.; Wang, D. X.; Yu, Y.; Yang, H. B. Stimuli-Responsive Supramolecular Gels through Hierarchical Self-Assembly of Discrete Rhomboidal Metallacycles (vol 19, pg 10094, 2013). *Chem. Eur. J.* **2013**, *19*, 12923-12923.

171. Singh, R.; Zhou, H. Mass tags for quantitative analysis **2002**, *US2002164649 (A1)*.

172. Pamukcu, R.; Piazza, G. A. Method of treating a patient having precancerous lesions with phenyl quinazolinone derivatives *Cell Pathways Inc.* **2003**, *US6562830 (B1)*.

173. Wensbo, D.; Louise, E.; Methvin, I.; McLeod, D. A.; Slassi, A.; Xin, T.; Stormann, T. M. *Astrazeneca AB., NPS Pharma Inc.* **2005**, *WO2005066155 (A1)*.

174. Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J. Appl. Crystallogr.* **1993**, *26*, 795-800.
175. Bailey, S. The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **1994**, *50*, 760-763.
176. Mccoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658-674.
177. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2002**, *58*, 1948-1954.
178. Saitou, N.; Nei, M. The Neighbor-Joining Method - a New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **1987**, *4*, 406-425.
179. Zuckerkandl, E.; Pauling, L. Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins* Academic Press: New York, 1965; pp 97-166.
180. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725-2729.
181. Velazquez-Campoy, A.; Freire, E. Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat. Protoc.* **2006**, *1*, 186-191.

# 4

# TOWARDS THE DISCOVERY OF SMALL MOLECULE ACTIN BINDERS

Actin is the most abundant eukaryotic protein and participates in many important cellular functions such as muscle contraction, cell motility, cell division, vesicle and organelle movement, cell signaling and establishment and maintenance of cell junctions and cell shape. The importance of the actin cytoskeleton in pathogenic processes such as angiogenesis and metastasis has made it an attractive target for the development of new anticancer drugs. However, up to date no small organic molecules have been developed that could validate the therapeutical potential of actin. Instead, there are a variety of complex natural products that interact with this protein, whose therapeutical applications have not been studied due to the scarce amounts in which these molecules can be isolated from their natural sources.

The project presented in this chapter aims, in collaboration with the group of Prof. Amedeo Caflisch (Biochemistry Department, UZH), to apply and develop new methods for the *in silico* discovery of novel actin binding molecules inspired by natural product "small molecules" (NPs).

In the next section, a general overview of the actin cytoskeleton and the known actin binding natural products is given. To the introduction section, the results and discussion follow, and the chapter finishes with a conclusion section.

## 4.1 Introduction

### 4.1.1 Cytoskeleton fibers

The cytoskeleton is a dynamic three dimensional structure that fills the cytoplasm in eukaryotic cells and acts as both, motor and framework.[1] The cytoskeleton is mostly

composed of long fibers (polymers) stemming from the repetitive assembly and disassembly of certain proteins.

Three different types of fibers can be distinguished within the cell cytoskeleton according to their composition and size (Figure 1). The smallest fibers (3-6 nm in diameter) are referred to as microfilaments and are composed of actin (Figure 1, left), the protein of interest in this work. The biggest fibers (20-25 nm in diameter) are called microtubules, they adopt the shape of hollow cylindrical tubes and are composed of the cytoskeletal protein tubulin (Figure 1, middle). The size of intermediate filaments lies between microfilaments and microtubules, as they are 10 nm in diameter (Figure 1, right). Intermediate filaments are heterogeneous structures composed of a variety of fibrous proteins that provide the tensile strength to the cell and can only be found in organisms with multicellular organization such as vertebrates, nematodes and mollusks.

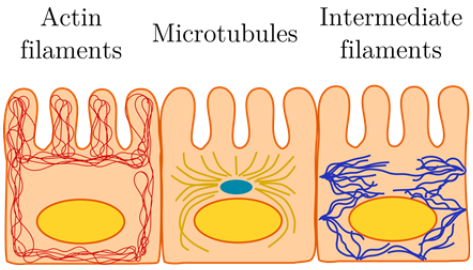



**Figure 1.** Components of the cytoskeleton in the epithelial cells of the intestine: actin filaments (red), mictrotubules (yellow) and intermediate filaments (blue). Microfilaments give the shape to the cell surface, while microtubules grow out of the centrosome to the cell periphery and intermediate filaments connect adjacent cells.[2]

Out of the three above-mentioned types of cytoskeletal fibers, tubulin has been the most studied by far. Tubulin is a 50 kDa in size globular protein that represents 3-4 % of the total protein content in all eukaryotic cells. It can be found in different forms (α-, β-, γ-, δ-, ε-, and ζ-tubulin) depending on its cellular location and function. α- and β-Tubulin subunits form heterodimers, which assemble into protofilaments that build up microtubules of variable length (200 nm – 25 µm) as depicted in Figure 2. For heterodimers to associate between each other forming protofilaments, both, α- and β-tubulin subunits must have a GTP (guanosine triphosphate) molecule bound. As the heterodimers associate, the GTP bound to β-tubulin hydrolyses into GDP (guanosine diphosphate), which remains bound to the α-tubulin subunit (close-up in Figure 2). Finally, the protofilaments associate laterally to form sheets, and eventually microtubules are assembled. During this process, accessory proteins are able to influence the stability and dynamics of mictrotubule assembly.

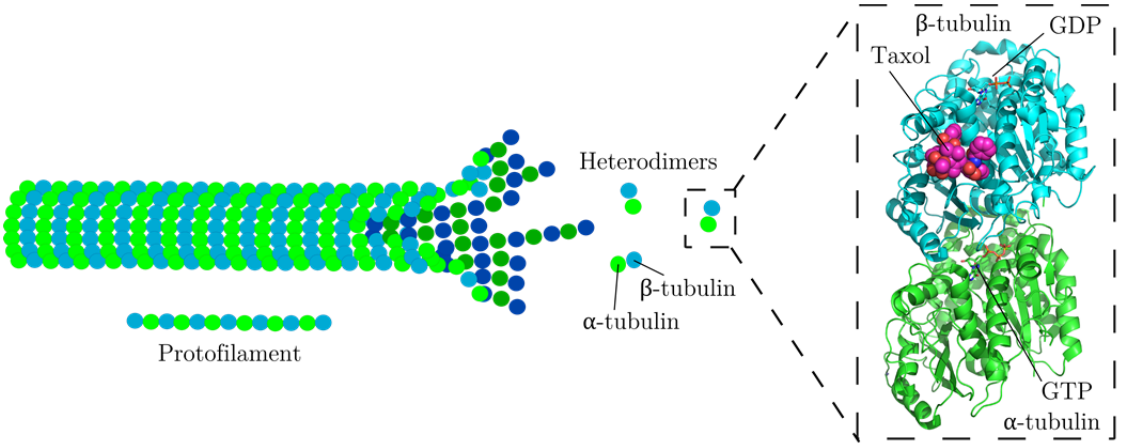



**Figure 2.** Structure of microtubules. Microtubules are formed by the macroassembly of protofilaments into hollow cylinders. The protofilaments are composed of repeating units of α- and β-tubulin heterodimers. A tubulin heterodimer in complex with the tubulin stabilizer taxol (shown as a surface representation in magenta) is given as a close-up on the right hand side. PDB code 1JFF.[2]

Mircrotubules are responsible for a wide variety of key cellular functions. They provide the structural framework that maintains the cell shape and influences the organization of various organelles and components of the cytoskeleton, as well as the formation of an internal transport network for vesicles and organelles (including mitochondria and lysosomes). Moreover, the rigid internal core provided by microtubules is employed by microtubule associated motor proteins to generate force and movement in motile structures such as cilia and flagella, that are tail like protusions of certain prokaryotic and eukaryotic cells that play an important role in cell locomotion and as sense organs. More importantly, microtubules form the spindle apparatus during cell division (Figure 3) ensuring the proper segregation of duplicated chromosomes into daughter cells during cytokinesis, the process in which the cytoplasm of a cell is divided into two. The spindle apparatus also regulates the assembly and location of an actin-rich contractile ring that pinches and separates the two daughter cells (see section 4.1.2 for further details).

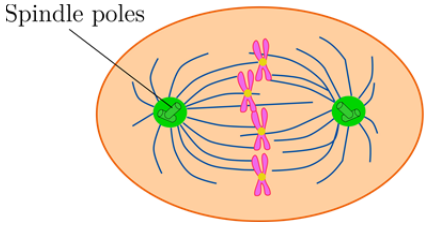



**Figure 3.** Spindle apparatus formed during cell division.

Due to its key biological functions, tubulin has been widely studied and several compounds able to interfere with tubulin assembly have been reported and applied for cancer treatment, such as paclitaxel, also known as taxol (Figure 2).[3] In contrast, actin has recently been recognized as a potential therapeutic target and its key role in numerous diseases is starting to be unraveled. A general introduction to actin focusing on actin binding proteins and natural products is given in the next sections.

### 4.1.2 Actin structure: G actin vs. F actin

Actin is the most abundant protein in eukaryotic cells representing 5-10 % of the total protein content in cells, increasing up to 20 % in muscle.[4] Actin is also one of the most highly-conserved proteins in nature, sharing more than 80 % of its structural constituents in species as remotely related as humans and algae.[5]

Actin exists in monomeric form (referred to as globular or G-actin) or as a polymer (known as filamentous or F-actin). G-actin is 43 kDa in size and can be divided into four subdomains (Figure 4, right).[6, 7] The so called cleft region is located in the middle of the four subdomains and is usually occupied by an ATP (adenosine triphosphate) or ADP (adenosine diphosphate) molecule, together with a divalent cation, $Ca^{2+}$ or $Mg^{2+}$, (Figure 4, right).[8] G-actin can reversibly self-assemble through non-covalent interactions into long and double-stranded helical polymers known as filamentous or F-actin (Figure 4, left and middle). Unfortunately, no high-resolution X-ray structure of F-actin is yet available, and thus, several models of the filament have been proposed.[9, 10]

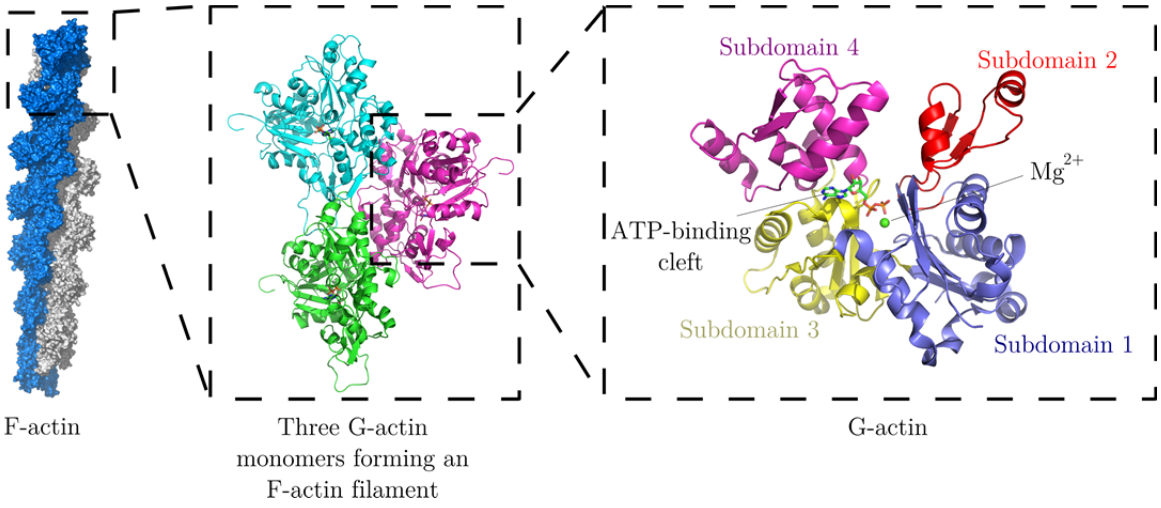



**Figure 4.** Left: Representation of F-actin derived from Holme's model.[9] Middle: Close-up of a model of three actin monomers forming an F-actin filament.[10] Right: X-ray structure of G-actin in complex with ATP (shown as sticks) and a $Mg^{2+}$ ion (PDB code 1WUA). The four subdomains are indicated in different colors and the ATP-binding cleft is specified.

### I. Mechanism of actin polymerization

The ATP-dependent polymerization mechanism in which G-actin assembles into F-actin can be divided into four reversible steps: activation, nucleation, elongation and annealing (Figure 5).[11]

As a first step of the actin polymerization process, salts induce a conformational change in G-actin, causing ATP-bound monomers (three or more) to associate with each other forming oligomers (referred to as the nucleation step). This step constitutes the rate determining step of actin polymerization, which generates a lag time.[5]

After the nucleation sites are formed, the elongation step starts: F-actin polymers grow at both ends of the filament, ‘‘pointed end’’ (-) and ‘‘barbed end’’ (+), generating filament polarity. The “barbed end” or “plus end” (which elongates five to ten times faster) is comprised of subdomains 2 and 4, whereas the “pointed end” or “minus end” consists of subdomains 1 and 3 (Figure 5, close-up).[12, 13] During polymerization, ATP hydrolyses into inorganic phosphate and ADP, resulting in a conformational change in subdomain 2 of actin (Figure 5, close-up).[8, 14] Elongation continues until the monomer concentration reaches the critical concentration (see Figure 6 for further information).

Dissociation of actin monomers from the filament takes place mainly at the “pointed end”).[11] At the end of the polymerization, the actin filament assembly reaches a steady state where dissociation and association are equal and the actin filaments appear to “treadmill”, maintaining the filament length almost constant.[15] As a final step, actin filaments can also associate with each other in a process called annealing.[16]

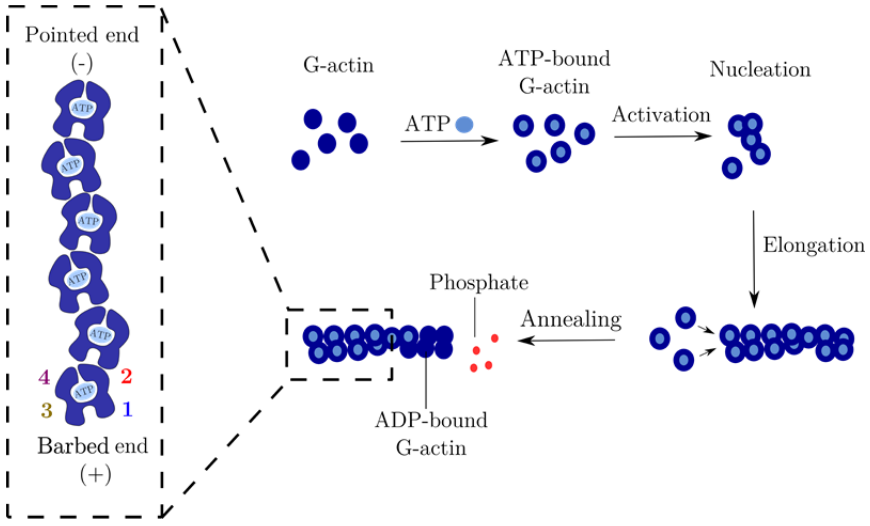
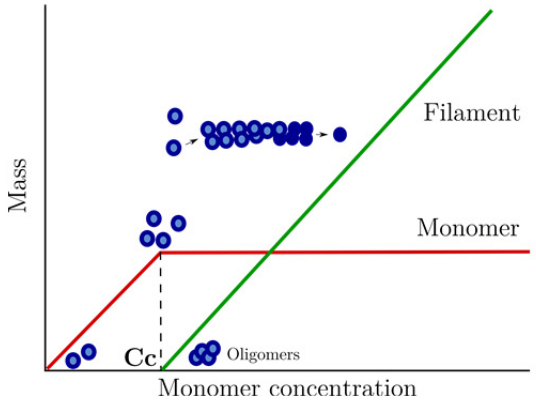



**Figure 5.** ATP-dependent polymerization of G-actin into F-actin. As a close-up a schematic representation of F-actin is given, where the four different subdomains and the two filament ends, the ‘‘barbed end’’ (or + end) and the ‘‘pointed end’’ (or – end), are indicated.

In order for polymerization to take place, several factors such as pH and salt concentration need to fall under a controlled range. Furthermore, even at optimal pH and salt concentration, a minimum concentration of G-actin monomers is required, referred to as the critical concentration (Cc). The critical concentration (Figure 6) is defined as the monomer concentration at which the association rate (rate of G-actin addition onto the filament, $k_+$) equals the dissociation rate (rate of G-actin dissociation from the filament, $k_-$). Dissociation is dominant below the critical concentration, whereas above the critical concentration the association process is favored. An estimation of actin association and dissociation rates on actin is given in Figure 6, right.




| | Cc | $k_+$ | $k_-$ |
|---|---|---|---|
| **(-) end** | 0.6 µM | 1.3 µM $^{-1}$ s $^{-1}$ | 0.8 s $^{-1}$ |
| **(+) end** | 0.12 µM | 12 µM $^{-1}$ s $^{-1}$ | 1.4 s $^{-1}$ |

**Figure 6.** Left: Schematic representation of the effect of the critical concentration (Cc) during the actin filament assembly process. Actin filaments start polymerizing only at concentrations above the Cc.[2] Right: Actin association and dissociation rates on actin filaments. Relative actin association ($k_+$) and dissociation rates ($k_-$) depend on the monomer concentration and the critical concentration (Cc).[17]

The value of the critical concentration depends on several conditions such as ionic strength of the solution, presence of divalent cations, pH and temperature. At low salt concentrations, for instance, the critical concentration is high and actin will not polymerize even at high monomer concentrations. Alternatively, the addition of salt, 20-100 mM KC1 and/or 0.5-5 mM $MgCl_2$, to a solution of monomers induces polymerization.[18] The effect that the pH, ionic strength, nucleotide type, ions and temperature have on actin polymerization has been summarized in Table 1.

**Table 1.** Effect of solution conditions on actin polymerization.[19]

| | **Favours F-actin** | **Favours G-actin** |
|---|---|---|
| pH | Low (5-6) | High (8-9) |
| Ionic strength | Moderate (0.1M NaCl) | Low (zero) |
| Nucleotide | ATP | ADP |
| Divalent cation | $Mg^{2+}$ | $Ca^{2+}$ |
| Temperature | High (37 °C) | Low (4 °C) |

### II. Biological function of actin

The actin cytoskeleton is responsible for a wide range of key cellular functions. Due to its strong and dynamic character, it guarantees cell durability, maintains the cell shape and anchors cytoskeletal proteins; whereas thanks to its flexibility it can accommodate the multiple and varied demands of the cell.[20-23]

Actin microfilaments interact with other cytoskeleton fibers (see section 4.1.1). On the one hand, they cooperate with microtubules via microtubule-associated proteins during the transport of vesicles and organelles.[24] On the other hand, they enable extracellular stimuli to be transmitted to key targets like ribosomes and chromosomes in cooperation with intermediate filaments.[25]

The distribution of actin filaments is highly dependent on the specific function of the cell or tissue. In polarized cells such as those of the epithelium, actin filaments play a key role in tissue organization and establishment of cell polarity and cohesion (Figure 7). As an example, a dense bundle of cross-linked actin filaments form the structural core of microvilli, which are microscopic membrane protrusions that increase the surface area of cells (Figure 7, left). Furthermore, the integrity of epithelial cell layers is maintained by a belt of actin (adhesion belt, Figure 7, middle), which links the cytoskeleton between adjacent cells, and by a cell cortex (Figure 7, right), which determines the shape, stiffness and movement of the cell surface.

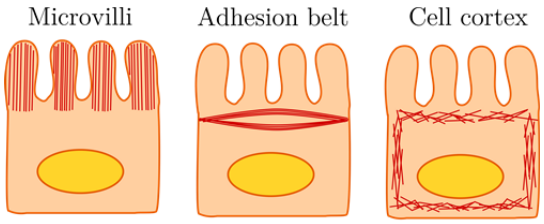



**Figure 7.** Actin filament distribution in epithelial cells.

Actin filaments are the primary cytoskeleton components that drive cell motility (Figure 8), and as such, they are highly abundant in membrane protrusions that are used by cells to migrate (called filopodia and lamellipodia). Cell movement is essential for the formation of tissue during embryonic development, wound healing and immune responses, and errors during this process could have serious consequences, including mental retardation, vascular disease, tumor formation and metastasis.

Two main theories have been developed to explain the movement of a cell's front edge: the cytoskeletal and the membrane flow model. According to the cytoskeletal model (Figure 8, A) the cell cortex (Figure 7, right) in moving cells is concentrated towards the

leading edge, where rapid actin polymerization occurs.[26] The formation of actin filaments pushes the leading front forward and, as a consequence, the cell migrates.[27, 28] On the other hand, the membrane flow model states that the extension of the leading edge is the result of membrane addition through exocytosis (Figure 8, B).[29]

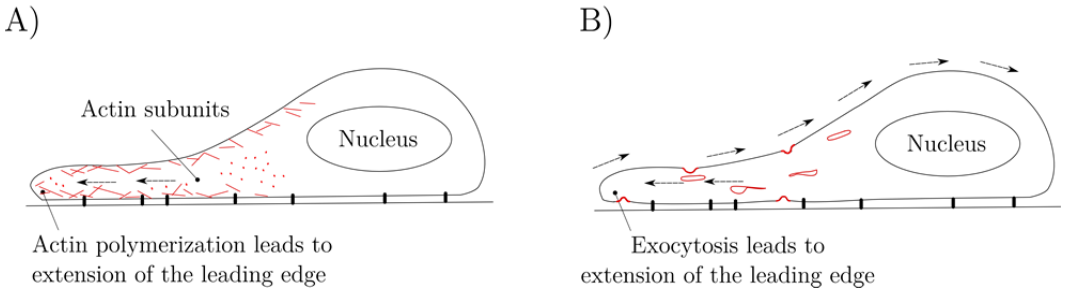



**Figure 8.** Two main models of cell migration. A) Cytoskeletal model: actin polymerization drives cell movement. B) Membrane flow model: membrane is brought towards the leading edge of the cell through exocytosis.

In muscle cells, actin-associated myosin motor proteins use ATP hydrolysis that takes place upon actin polymerization to apply forces against stress fibers during muscle contraction.[30] The energy liberated during ATP hydrolysis can also be converted to tensile force that retracts the edge of motile cells. Alternatively, in non-muscle cells, actin filaments function as "track" for cargo transport myosins (non-conventional myosins), which use ATP hydrolysis to transport vesicles and organelles.

The actin-based structures previously described are disassembled before cell division, which causes the cell to stop moving and become rounded. During cell division, actin bundles play the key role of orienting the microtubule network. Moreover, the proper assembly, orientation, and contraction of the actin filament ring (contractile ring) pinches and separates the daughter cells during cytokinesis in cell division (Figure 9).

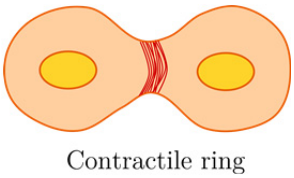



**Figure 9.** Actin filament distribution during cytokinesis in cell division.

### 4.1.3 Actin binding molecules

#### I. Actin binding proteins

As seen in the last section, the actin cytoskeleton is of upmost importance for a wide variety of cellular processes that demand an accurate control of the interchange

between monomeric G-actin and polymeric F-actin filaments. Proteins that bind to G-actin monomers or F-actin filaments are able to influence the actin filament assembly and disassembly process in different ways that are classified as follows: monomer ‘‘sequestering’’ proteins, nucleating proteins, ‘‘capping’’ and ‘‘severing’’ proteins, stabilizing and crosslinking proteins (Figure 10).[31, 32]

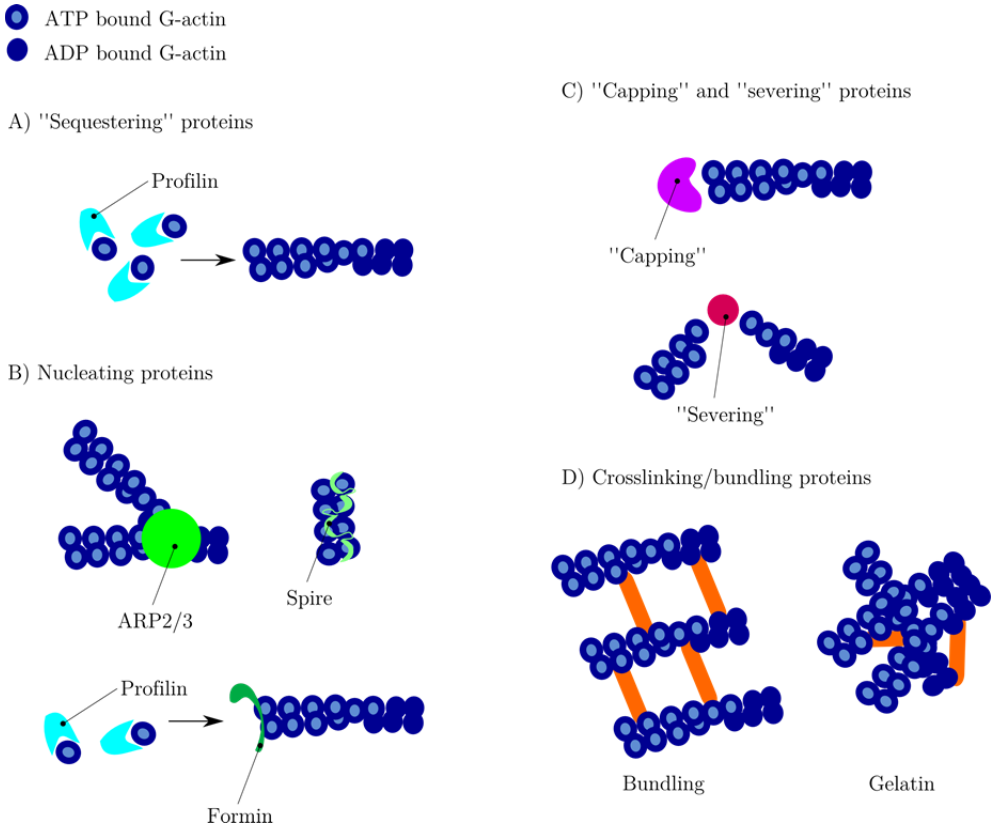



**Figure 10.** Overview of actin binding proteins classified according to their mode of action in ‘‘sequestering’’, nucleating, ‘‘capping’’, ‘‘severing’’, stabilizing (not shown) and crosslinking proteins.[32]

G-actin monomer binding proteins (also called “sequestering” proteins) are able to bind to G-actin monomers. Profilin is a representative example, as it facilitates the exchange of ADP by ATP in G-actin, thus enabling actin polymerization (Figure 10, A).[33]

Actin nucleating proteins ease oligomer formation during nucleation, which is the rate determining step of the actin polymerization process. The ARP2/3 complex is a nucleating protein that binds to two G-actin monomers and the side of an F-actin filament simultaneously, resulting in the formation of a new F-actin filament (Figure 10, B).[28] Another example of nucleating proteins are formins, which in cooperation with profiling (a “sequestering” protein) nucleate new filaments.[34] Spire has also been recently recognized as an actin nucleator that is able to recruit four actin monomers forming a nucleation seed for actin polymerization to begin.[35]

The so called "capping" proteins cover one of the F-actin filament ends, inhibiting as a consequence, the addition and loss of G-actin monomers. "Severing" proteins, on the other hand, promote the disassembly of actin filaments. CapZ, for instance, prevents polymerization by "capping" the "barbed end" of the filament, whereas cofilin is able to "sever" actin filaments inducing "pointed end" depolymerization.[36] Interestingly, proteins like gelsolin are able to perform both processes, "cap" and "sever" F-actin filaments.[37]

Another type of actin binding proteins are classified as crosslinking proteins, which influence the packing and organization of actin filaments into secondary structures, including unordered networks (filamin), tight (fimbrin) and loose ( $\alpha$ -actinin) bundles.[38]

The last type of actin binding proteins, the so called stabilizers, such as tropomyosin, stabilize F-actin filaments inhibiting their depolymerization.

Alternatively, actin filament assembly can be modulated by external processes (referred to as upstream regulation), such as controlled nucleotide hydrolysis (e.g. ATP on actin) and reversible modifications (e.g. phosphorylation) on components that control the actin assembly process. As an example, cofilins are inhibited as soon as they are phosphorylated by kinases.[36] As their phosphate group is removed by dephosphatases, cofilins are reactivated and are able to regulate the actin polymerization process by "severing" F-actin filaments.

### II. Actin binding natural products

In addition to the extensive research and the wide variety of proteins known to interact with actin, actin binding natural products have also greatly contributed to the study of actin dynamics. Most of these secondary metabolites are of marine origin and play a role in the chemical defense of the organisms where they originate. Interestingly, they are able to mimic the mode of action of proteins by "sequestering" monomers, "capping" and/or "severing" filaments and promoting polymerization by nucleation or by stabilization of the filaments. Actin binding natural products are classified into two main groups: destabilizers (which inhibit F-actin filament assembly or destabilize filaments) and stabilizers.

*a) Actin destabilizers*

There are two different binding pockets in G-actin where actin destabilizers are known to bind: the ATP binding site (only targeted by latrunculins) and the "barbed end" (Figure 5).

Most actin destabilizers interact in a hydrophobic region of the "barbed end" (such as swinholide A,[39] misakinolide,[40] reidispongiolide,[41] sphinxolide B,[42] mycalolides,[43]

kabiramides,[43] aplyronine,[44] bistramide[45, 46] and cytochalasins[47]). These molecules possess a large macrocycle and display a long and stereochemically complex side chain usually terminated in a highly polar functional group (e.g. formamide, Figure 11).

Their macrocyclic moiety binds in a hydrophobic cleft between subdomains 1 and 3 of the "barbed end", followed by the intercalation of the molecule's tail into the cleft separating the two subdomains.[38, 48] Three examples of natural products that bind within subdomains 1 and 3 and their binding mode are given in Figure 12.

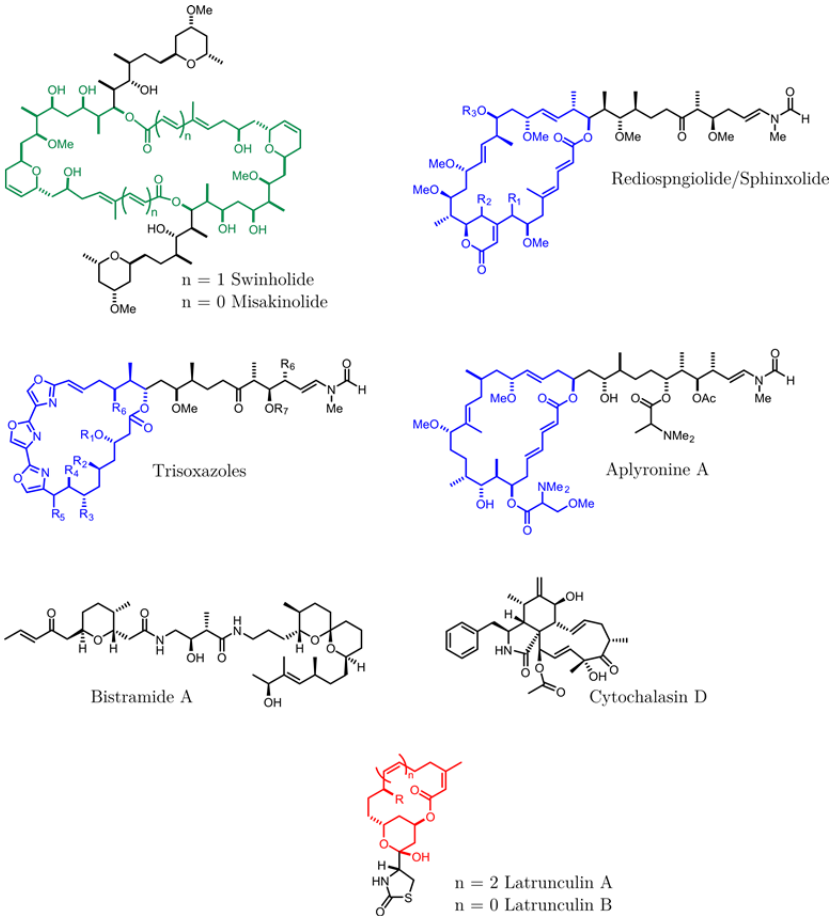



**Figure 11.** Structural diversity in actin destabilizing natural products of marine origin.


**Figure 12.** Structures of aplyronine A (PDB code 1WUA), bistramide A (PDB code 2FXU) and swinholide A (PDB code 1YXQ) in their respective complexes with actin. Actin is shown as a ribbon representation and toxins are shown as sticks in light blue. The toxins interact with the ‘‘barbed end’’ of the filament, between subdomains 1 and 3.

Taking their mode of action into account, actin destabilizing toxins targeting the “barbed end” can be divided into those that “sequester” actin monomers (either in a 1:1 or in a 1:2 ratio), “severe” F-actin filaments by disrupting interactions between adjacent monomers (such as mycalolides,[43] aplyronine A[44] and bistramide A[45, 46]) and/or “cap” the end of the filament inhibiting monomer association or dissociation (such as cytochalasins[47] and misakinolide[40]). Interestingly, several natural products are able to do both, “cap” and “sever” actin filaments (such as swinholide A,[39] reidispongiolide,[41] sphinxolide[42] and kabiramides[43]). An schematic description of the above-mentioned mode of actions is provided in Figure 13, and further details can be found in Table 2.

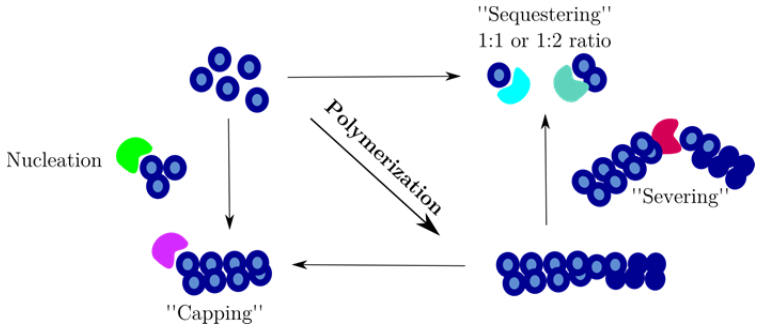



**Figure 13.** Possible mode of action of actin destabilizers. Actin destabilizers can “sequester” monomers and “cap” or “sever” the F-actin filaments.

Bistramide A and cytochalasin are two structurally different actin destabilizers that bind within the “barbed end” (Figure 12) that are worth mentioning.[45-47] The interaction of Bistramide A with actin takes place through the entire length of the hydrophobic cleft and it involves more polar contacts than other natural products (Figure 12, middle).[45, 46] Bistramide A does not only “sequester” actin in a 1:1 ratio, but it is the only actin destabilizer able to covalently modify actin.[45, 46] Cytochalsin, on the other hand, promotes

nucleation and "caps" the "barbed end" of the filament (Table 2), a unique mode of action among actin destablizers.[47]

**Table 2.** Binding site, mode of action and main structural features of actin destabilizers.

| Binding type | Natural product | Mode of action | | | | Structural features |
|---|---|---|---|---|---|---|
| | | Sequestering ratio | Severing | Capping barbed end | Other | |
| Bind at the nucleotide cleft | Latrunculins[49, 50] | 1:1 | - | - | - | Medium ring size. Thiazolidinone ring, no side chain |
| Bind between subdomains 1-3 | Swinholide A[39] | 2:1 | ✓ | ✓ | - | > 40 membered ring macrolides, symmetric structures, 2 lateral chains |
| | Misakinolide[40] | 2:1 | - | ✓ | - | |
| | Reidispongiolide[41] | 1:1 | ✓ | ✓ | - | 25-30 membered ring macrolides, one polar lateral chain |
| | Sphinxolide B[42] | 1:1 | ✓ | ✓ | - | |
| | Mycalolide[43] (trisoxazole) | 1:1 | ✓ | - | - | |
| | Kabiramide[43] (trisoxazole) | 1:1 | ✓ | ✓ | - | |
| | Aplyronine[44] | 1:1 | ✓ | - | - | |
| | Bistramide[45, 46] | 1:1 | ✓ | - | Covalent modification of actin | Polyether |
| | Cytochalasins[47] | 1:1 | - | ✓ | Promotes nucleation | Rigid bicycic isoindolone core fused to a macrocycle |

Latrunculins, which are the focus of this work, are the only actin binding natural products able to bind within the ATP binding cleft that is located at the "pointed end" of the actin filament. Latrunculin cytotoxins are isolated from the Red Sea sponge *Negombata Magnifica*[49, 50] and are very efficient actin polymerization inhibitors by "sequestering" G-actin in a 1:1 complex.[51] They possess a biologically rare thiazolidinone ring that is buried in the binding site and a hydrophobic macrocyclic ring. The thiazolidinone ring is involved in five intramolecular hydrogen bonds with the protein (one of which is water mediated) as illustrated in Figure 14.

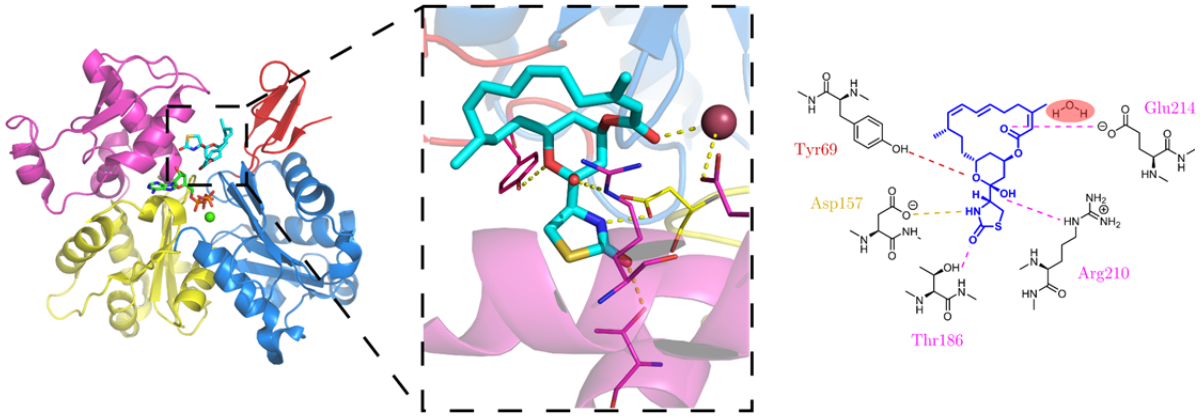



**Figure 14.** Crystal structure of actin-latrunculin complex, PDB code 1ESV.[49] As a close-up, the hydrogen bonds that latrunculin forms with G-actin are shown as dashed lines, in a 3D representation. On the right, a 2D representation of latrunculin's binding mode is given. Latrunculin is shown in blue and the hydrogen bonds are shown as dashed lines

*b) Actin stabilizers*

Actin stabilizers constitute another class of actin binding natural products. No X-ray structure of actin stabilizers with F-actin has yet been solved in complex with actin. However, information about the binding mode of actin stabilizers has been been studied based on X-ray fiber diffraction data, electron microscopy and biochemical analysis.[52-55]

Actin stabilizers present a cyclic depsipeptide as common structural feature (Figure 15). As well as stabilizing F-actin filaments, some actin stabilizers are able to promote nucleation and/or induce actin polymerization in non-polymerizing conditions (Table 1), referred to as nucleators (such as jasplakinolide,[56] chondramides[57] and dolastatin[58]). A more detailed explanation of actin stabilizers' mode of action is given in Table 3.

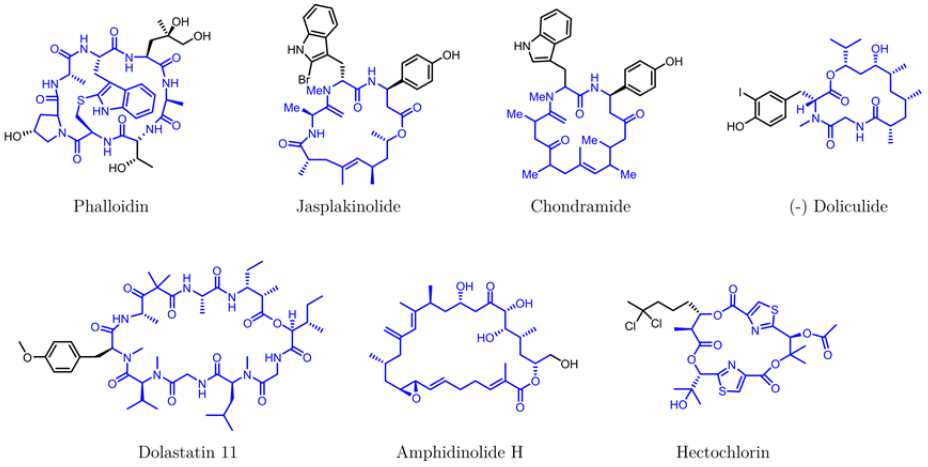



**Figure 15.** Chemical structure of some actin stabilizing natural products.

**Table 3.** Mode of action and main structural features of actin stabilizers.

| Natural product | Mode of action | | | | | Structural features |
|---|---|---|---|---|---|---|
| | F-actin stabilizer | Polym. inducer[a] | Nucleation promoter[b] | Competition with phalloidin binding site | Other | |
| Phalloidin[59] | ✓ | - | - | - | - | Bicyclic heptapeptide |
| Jasplakinolide[56] | ✓ | ✓ | ✓ | ✓ | - | 16-19 membered cyclic depsipeptide |
| Chondramide[57] | ✓ | ✓ | n.r. | ✓ | - | |
| Doliculide[60] | ✓ | ✓ | n.r. | ✓ | - | |
| Dolastatin[58] | ✓ | ✓ | n.r. | - | - | 30 membered cyclic depsipeptide |
| Amphidinolide H[55, 61] | ✓ | - | - | - | Covalent modification of actin | 26 membered macrolide. Allyl epoxide |
| Hectochlorin[62] | ✓ | ✓ | ✓ | - | - | 19 membered macrolyde |

[a]Induces actin polymerization under non polymerizing conditions. [b]Helps the nucleation (rate determining step), thereby decreasing the lag-time. n.r. not reported

Phalloidin is one of the best known actin filament stabilizing compounds.[59] Its binding mode has been characterized by microscopic and X-ray fiber diffraction methods, which showed that it interacts with three actin monomers simultaneously via three different binding sites enhancing filament stability (Figure 16).[53] Moreover, phalloidin has been labelled with different fluorophores, which have been successfully applied to visualize F-actin in cell based studies.

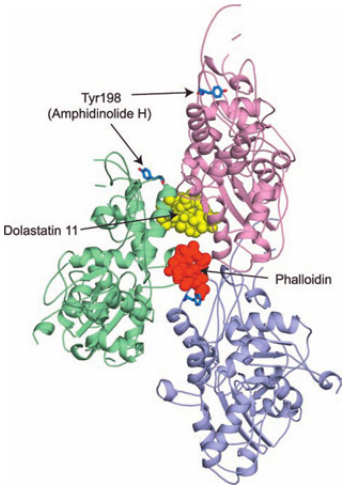



**Figure 16.** Binding sites of actin-stabilizing macrolides. Phalloidin interacts with three monomers simultaneously,[53] whereas dolastatin binds between subdomain 4 and subdomain 1 of the diagonal subunit.[63] Amphidinolide H (not shown) is known to form a covalent bond with Tyr198 in yeast or Tyr200 in mammalian actin (blue sticks).[55]

Amphidinolide H, presents a unique chemical structure among actin stabilizers with the presence of an epoxide ring.[55, 61] A tyrosine residue located in subdomain 4 of actin (Tyr198 in yeast, Tyr200 in mammalian actin) is able to nucleophilically attack the epoxide moiety forming a covalent bond.

## 4.2 *In Silico* design of actin inhibitors

### 4.2.1 Challenges and overview of the project

The discovery of new, versatile actin binding small molecules is highly valuable. First of all, actin binding molecules present a different and unique mode of action when interfering with the highly complex dynamics of the actin cytoskeleton, and they are thus valuable tools to study cellular pathways that are dependent on the actin cytoskeleton. Second, they are promising leads for anticancer drug development due to the direct implication of actin in cell motility (including cell invasion, division and metastasis)[12] and the high cytotoxicity displayed by some actin binding natural products.[21, 23, 64]

Cancer cells, which are defined as cells that have lost their regulatory mechanism to control their growth and proliferation, present an altered molecular composition, organization and polymeric state of the actin cytoeskeleton.[65, 66] As actin is present in all cells, the main limitation of the above-mentioned actin binding molecules is the potential inespecificity of clinical candidates to distinguish between healthy and malignant cells. However, this caveat is also present in tubulin targeting molecules such as taxol, which is used as an anti-cancer drug.[67] Nevertheless, as the rate of actin polymerization is enhanced in highly proliferative cells, the effect of actin binding molecules in cancer cells might be intensified, resulting in a cytostatic rather than in a cytotoxic effect.

The first goal of this project was to develop an *in silico* fragment-based docking approach inspired by natural product small molecules (NPs) for the discovery of inhibitors targeting the nucleotide cleft of G-actin. The structural information available from X-Ray structures of natural products interacting with G-actin has revealed two distinct binding sites: a hydrophobic cleft between subdomains 1 and 3 where molecules bind to monomeric or dimeric G-actin (e.g. swinholide[39], mycalolide[43], etc…) and a nucleotide cleft located between subdomains 2 and 4 (targeted by latrunculins).[51] We have chosen the nucleotide cleft as our working model due to the presence of polar interactions, especially buried hydrogen bonds that make the discovery of high affinity inhibitors more probable. The cleft located between subdomains 1 and 3, on the other hand, represents a shallower environment, and thus, the main contribution of inhibitor binding will take place via weak hydrophobic interactions.

An overview of the work-flow of the project is given in Figure 17. First, we aim to design a focused library based on natural products and ATP binding drugs that can offer a high specificity towards finding actin binders. The library will then be fragmented in a retrosynthetic manner and the obtained fragments will be docked, ranked and evaluated according to their binding free energy. Finally, in order to select the most promising fragments, molecular dynamic (MD) simulations will be performed. The selected fragments will then be synthesized and their affinity towards actin will be evaluated using *in vitro* and cell based techniques.

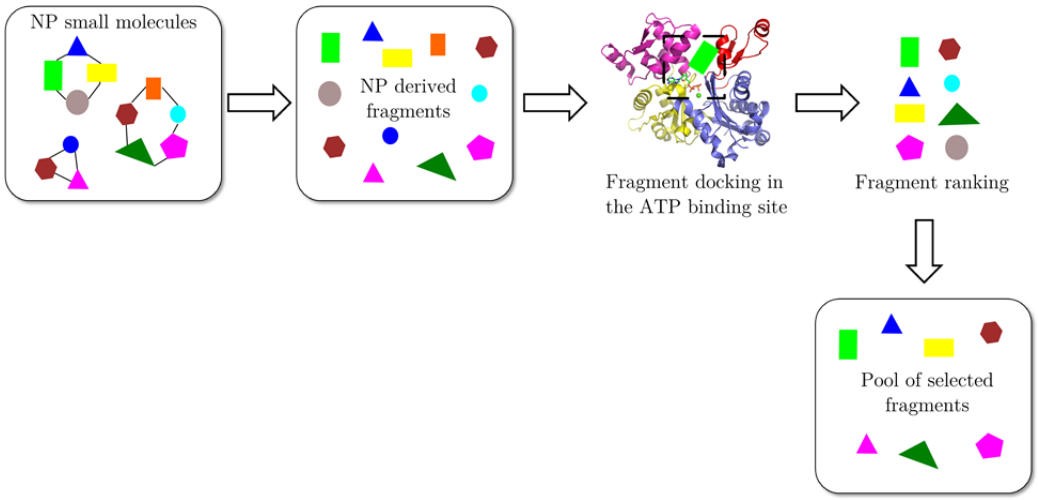



**Figure 17.** Work-flow of the project. Starting from a library of NP molecules, different fragments will be obtained and docked into the ATP binding site. The fragments will be ranked according to their predicted binding energy and selected for further studies, including chemical synthesis and biological evaluation.

### 4.2.2 *In silico* design

A general overview of the performed computer based design is given in this section. For further details, please refer to Sandra Rennebaum's PhD thesis in the group of Prof. Amedeo Caflisch (''Actin: computational design of novel inhibitors and study of the influence of the natural product latrunculin on conformational dynamics'', Biochemistry Institute, UZH, 2012).

First, a natural product based library composed of 140 000 compounds was constructed. Natural products are a major source of drugs, as more than 50% of antibacterials and 60% of anticancer compounds are based on nature's secondary metabolites.[68] Moreover, natural products are biologically pre-validated, relevant to nature and they can be considered as selected fractions of the chemical space during the course of evolution.[69] Furthermore, they offer a structural diversity that cannot be found in synthetically prepared chemical compounds, which enables the study of a new part of the

chemical space not provided by currently available libraries such as ZINC,[70] a commonly used library in *in silico* screening campaigns that is composed of commercially available molecules.

The natural product based library was compiled by combining natural product subsets from a variety of sources: natural product subset of the August 2009 version of the ZINC library, the Human Metabolome Database, Mircrosource Discovery and the Traditional Chinese Medicine Database. After the removal of compounds containing metals, a library composed of 140 000 molecules was obtained (Figure 18).

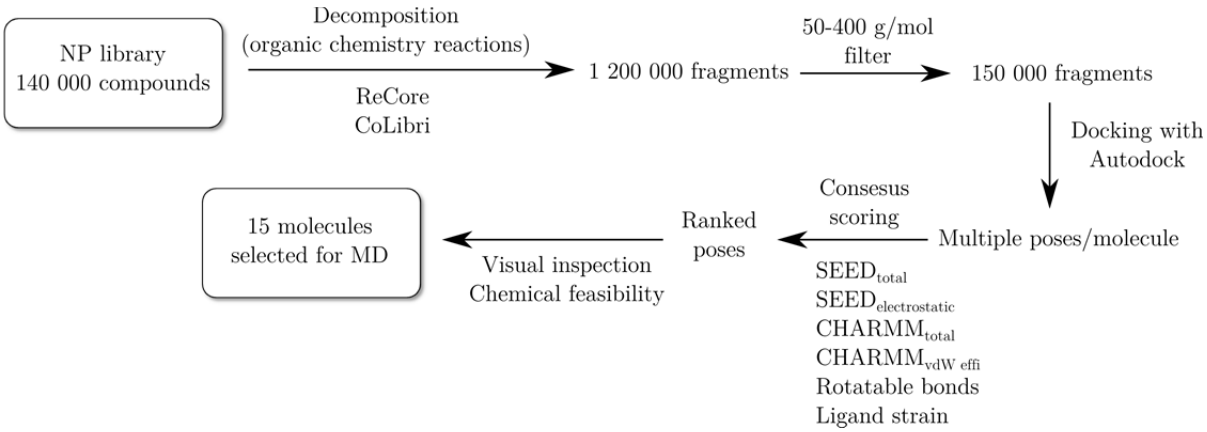



**Figure 18.** Work-flow of the *in silico* design of actin inhibitors.

The NP library was then decomposed in a retrosynthetic fashion according to specific organic chemical reactions using the software ReCore and CoLibri[71, 72] (Figure 19), so that all possible fragment combinations were taken into account in the next steps of the *in silco* design. As a result, a total of 1 200 000 fragments were obtained, which are chemically more feasible to synthesize.

Only fragments with a molecular weight between 50-400 g/mol and up to ten rotatable bonds were kept, resulting in 150 000 fragments. A 400 g/mol cut-off was applied in order to retrieve small fragments for a later lead optimization, which generally results in an increased molecular weight and worse drug-like properties. Those fragments were docked using AutoDock[73] and several poses per fragment (an average of 17) were obtained. An energy minimization of the poses in the rigid protein was then conducted using CHARMM.[79, 80]

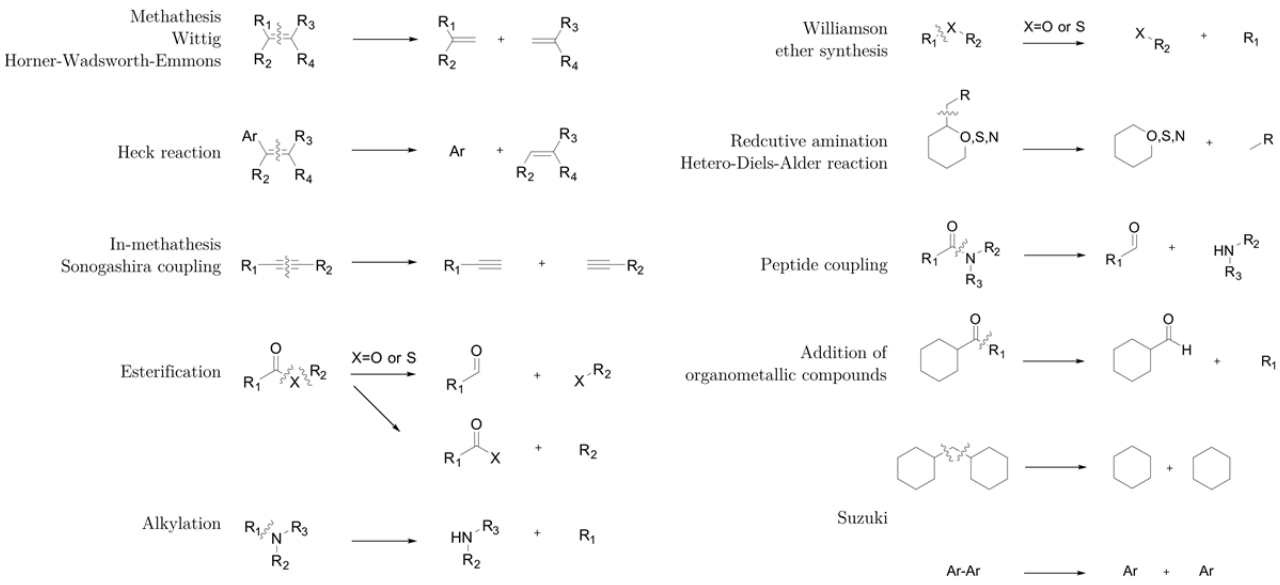



**Figure 19.** Cleavage rules designed according to organic chemical reactions.

In the next step of the *in silico* campaign an approach called consensus scoring was implemented. Consensus scoring combines several scoring functions that rank a set of docked poses according to their interaction energies, and has been shown to increase performance hit rates in virtual screening.[74, 75] Five consensus scorings involving different force fields and parameters were evaluated in order to select the most promising fragments (Figure 20).

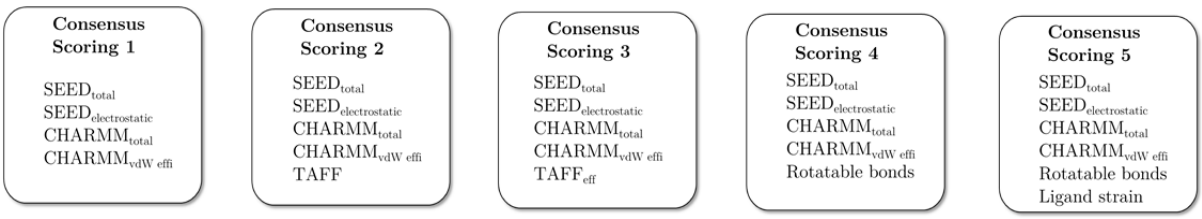



**Figure 20.** Overview of the implemented consensus scorings.

Different force fields were used in order to evaluate the interaction energies, including SEED (Solvation Energy for Exhaustive Docking),[77, 78] CHARMM[79, 80] (Chemistry at HARvard Molecular Mechanics) and TAFF (Tripos Atomic Force Field).[76] Both the SEED total interaction energy[77, 78] and the CHARMM total interaction energy[79, 80] account for electrostatic and van der Waals interactions. Moreover, the SEED total energy[77, 78] takes the desolvation of both the ligand and the receptor into account. TAFF, on the other hand, does not include electrostatic interactions, but only van der Waals and bonded interactions.[76] The CHARMM van der Waals efficiency is the CHARMM van der Waals interaction energy divided by the molecular weight. It is important to take the molecular

weight into account, as the largest molecules tend to give the most favorable energies due to a greater number of intermolecular interactions that they are able to form.

Together with the mentioned scoring functions, the number of rotatable bonds and ligand strain were considered (Figure 20, consensus scorings 4 and 5). The former is a simple approximation to incorporate the entropic penalty of the ligand upon binding, as ligands with many rotatable bonds present a higher entropic penalty due to the loss of rotational degrees of freedom upon binding. Ligand strain, on the other hand, allows discarding those molecules that adopt a high energy conformation upon binding.

The molecules were sorted according to the median rank of the interaction energies in order to discard outliers.[81] The top 100 fragments from consensus scorings 1 and 4, the 300 top ranking fragments from consensus scoring 5 and the 17 fragments appearing in the top 100 in the five consensus scorings were visually inspected using the software Witnotp and Pymol. The selection of the fragments was based on several parameters such as hydrogen bond analysis to the key residues involved in hydrogen bonds in the actin-latrunculin complex (Tyr69, Asp157, Thr186 and Arg210, Figure 14), hydrophobic interactions, structural diversity of the fragments and drug-like properties (by fulfilling Lipinsky's rule of five).[82] After visual inspection, fifteen fragments were selected to be further studied in molecular dynamics (MD) simulations (Figure 21).

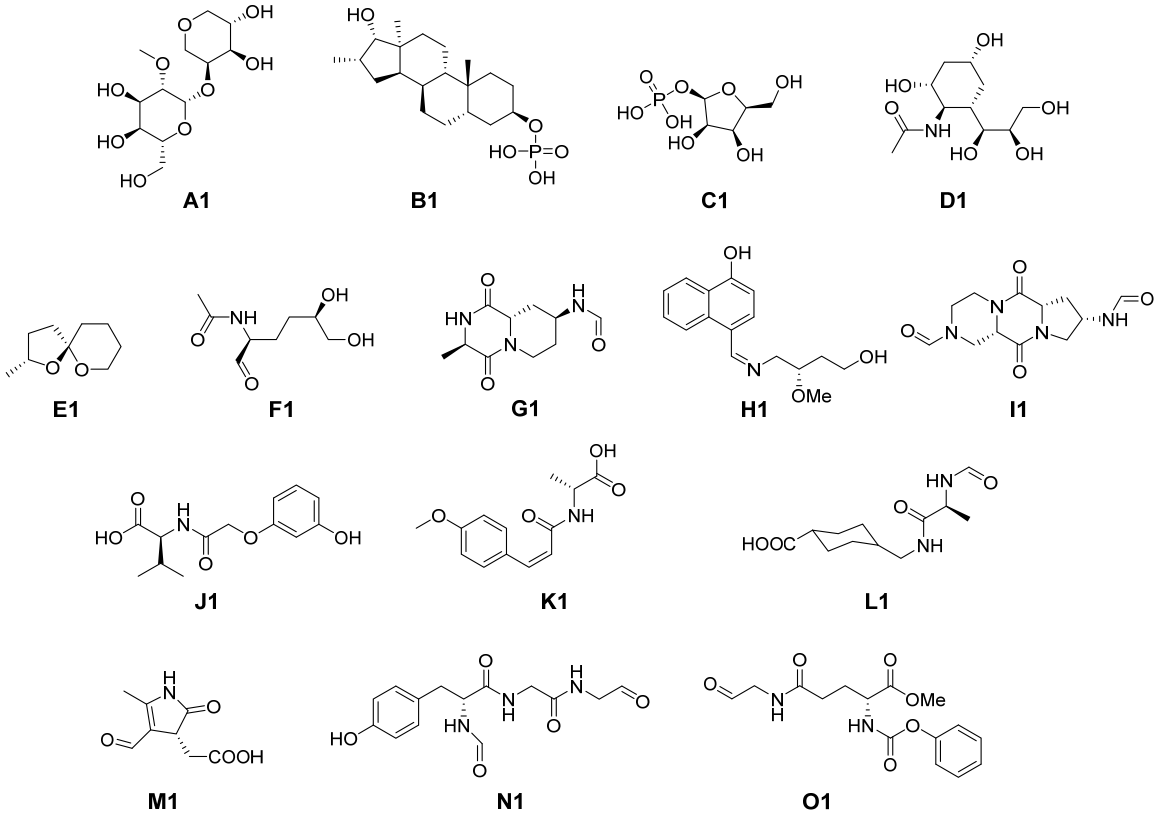



**Figure 21.** Selected fragments for MD simulations.

For each selected ligand 10-20 independent explicit water MD runs were performed starting from the docked pose obtained with AutoDock.[83] The potential affinity of the ligands was examined by taking the number of unbinding events into account (Table 4). An unbinding event was defined to take place if the separation of the center of mass of the ligand and the center of mass of the binding site was larger than 10 Å. Thus, a ligand that presents a low number of unbinding events is more likely to have a higher affinity for the protein.Time series of the distances between the ligands and the binding site indicated that most ligands present multiple binding modes and dissociation pathways.

**Table 4.** Overview of the MD simulations performed with compounds **A1**-**O1**.

| Compound | Number of runs | Number of unbinding events |
|---|---|---|
| **A1** | 15 | 1 |
| **B1** | 20 | 0 |
| **C1** | 15 | 0 |
| **D1** | 20 | 6 |
| **E1** | 10 | 9 |
| **F1** | 20 | 12 |
| **G1** | 20 | 14 |
| **H1** | 20 | 9 |
| **I1** | 20 | 6 |
| **J1** | 20 | 6 |
| **K1** | 16 | 7 |
| **L1** | 20 | 5 |
| **M1** | 15 | 3 |
| **enantiomer 1 M2** | 10 | 1 |
| **enantiomer 2 M2** | 10 | 0 |
| **N1** | 20 | 6 |
| **O1** | 20 | 7 |

This approach resulted in the selection of six scaffolds for synthesis. Compounds **A1**-**E1**, were automatically discarded for bad parametrization of the phosphate group (**B1**, **C1**), bad performance on the MD simulations (**E1**) or sugar character of the molecules (**A1**, **D1**). On the other hand, compounds **J1**-**O1** were selected due to their synthetic feasibility and performance on MD simulations (Table 4). Due to the high similarity of compound **M1** with the commercially available racemic mixture **M2**, MD simulations were run with both enantiomers (Table 4). Surprisingly, better results were obtained for compound **M2**, with 1 and 0 unbinding events in 10 independent runs. Thus, the biological evaluation was performed with the commercially available hit **M2**.

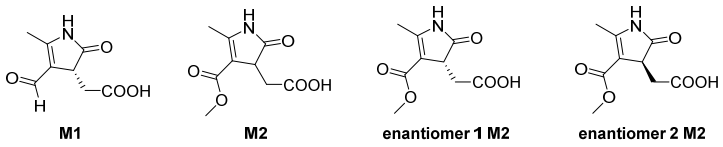


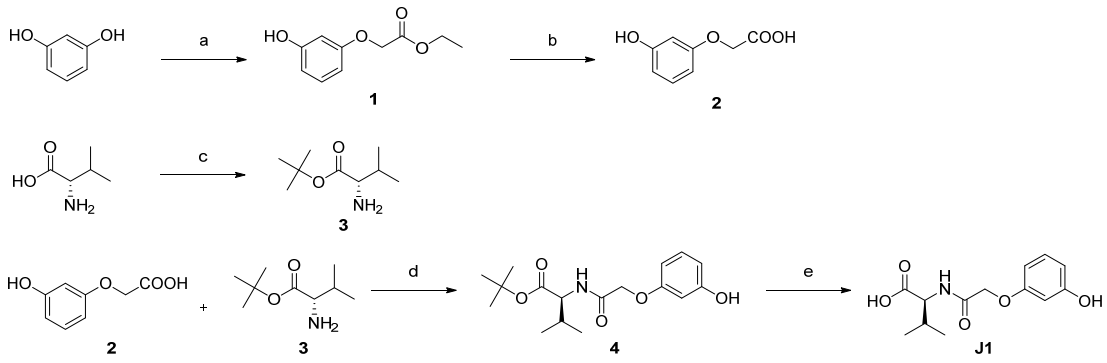

**Figure 21.** Comparison between compound **M1** and the commercially available compound **M2**

## 4.3 Chemical synthesis of selected molecules

### 4.3.1 Synthesis of compound J1

The synthesis of compound **J1** started from commercially available resorcinol, which was mono-etherified with ethylbromoacetate affording phenol **1** and subsequently hydrolyzed in the presence of NaOH into 3-hydroxyphenoxy acetic acid (**2**). Acid **2** was then coupled to L-valine *tert*-butyl ester (**3**) with DCC affording the desired amide **4** in moderate yield. A final hydrolysis of *tert*-butyl ester **4** provided the final product *N*-(3-hydroxyphenoxy)acetyl-L-valine **J1** in 52% yield (Scheme 1).

**Scheme 1**[a]



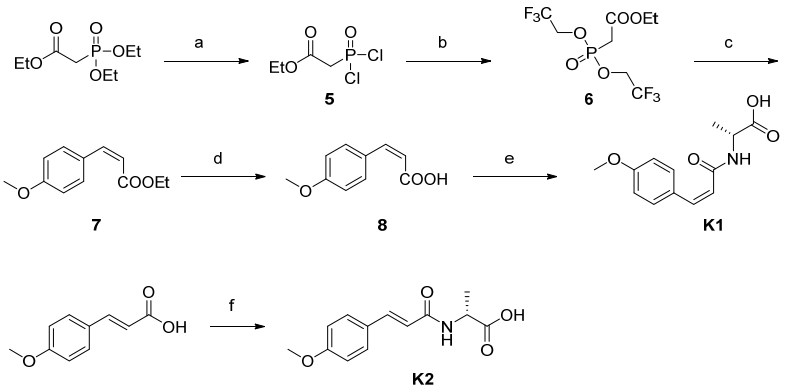

[a]Reagents and conditions: a) $BrCH_2CO_2Et$, $K_2CO_3$, DMF, 25 °C, 12 h, 30%; b) NaOH, EtOH, reflux, 12 h, 88% c) $^tBuOAc$, $HClO_4$, 25 °C, 12 h, 63%; d) DCC, HOBt, $Et_3N$, THF, 25 °C, 24 h, 64%; e) TFA, DCM, 25 °C, 4 h, 52%.

### 4.3.2 Synthesis of compounds K1 and K2

Triethyl phosphonoacetate was heated to reflux with phosphorus (V) pentachloride to afford **5**, which was reacted with 2,2,2-trifluoroethanol to obtain the bis(trifluoroethyl) phosphonester **6**. A Horner–Wadsworth–Emmons reaction under Still and Genari conditions was performed in the presence of *p*-anysaldehyde in order to obtain ethyl *cis*-4-

methoxycinnamate (**7**), which was subsequently hydrolyzed to *cis*-4-methoxycinnamic acid (**8**). As a final step, acid **8**) was coupled to D-alanine in the presence of DCC and *N*-hydroxysuccinimide affording *N*-(*cis*-*p*-methoxycinnamoyl)-D-alanine **K1** (Scheme 2).

**Scheme 2**[a]




[a]Reagents and conditions: a) $PCl_5$, reflux, 12 h; b) $F_3CCH_2OH$, $^iPr_2NEt$, benzene, 25 °C, 2 h, 49% over two steps; c) (i) $K[N(Si(Me_3)_2]$, 18-crown-6, toluene, -20 °C; (ii) *p*-anysaldehyde, 25 °C, 4 h, 47%; d) 1M NaOH, EtOH, reflux, 3 h, 92%; e) D-alanine, DCC, *N*-hydroxysuccinimide, dioxane, 25 °C, 12 h, 60%; f) D-alanine, DCC, *N*-hydroxysuccinimide, dioxane, 25 °C, 12 h, 63%.

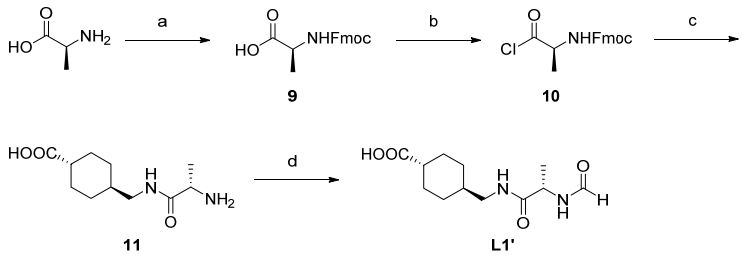
The synthesis of **K2** was carried out in an analogous manner by coupling the commercially available *trans*-*p*-methoxy cinnamic acid and D-alanine in the presence of DCC and N-hydroxisuccinimide (Scheme 2).

### 4.3.3 Synthesis of compound L1'

Instead of synthesizing molecule **L1** (Figure 21) with a *cis* relative configuration between the substituents in the cyclohexane ring, the *trans* stereoisomer was prepared due to the commercial availability of *trans*-4-(aminomethyl)cyclohexanecarboxylic acid. As the carboxylic acid is pointing towards the solvent in the predicted binding mode, the *trans* or *cis* disposition of the substituents will likely not have an influence on the molecule's activity. For simplicity, we will refer to this molecule as **L1'**.

The synthesis started by the Fmoc protection of L-alanine to give *N*-Fmoc-L-alanine (**9**). Upon treatment of **9** with $SOCl_2$, *N*-Fmoc-L-alanine acid chloride (**10**) was obtained, which was coupled to *trans*-4-(aminomethyl)cyclohexanecarboxylic acid under basic conditions to give **11**. As a final step, the free amine in **11** was reacted with acetic acid in the presence of acetic anhydride obtaining the corresponding formamide **L1'**.

**Scheme 3**[a]



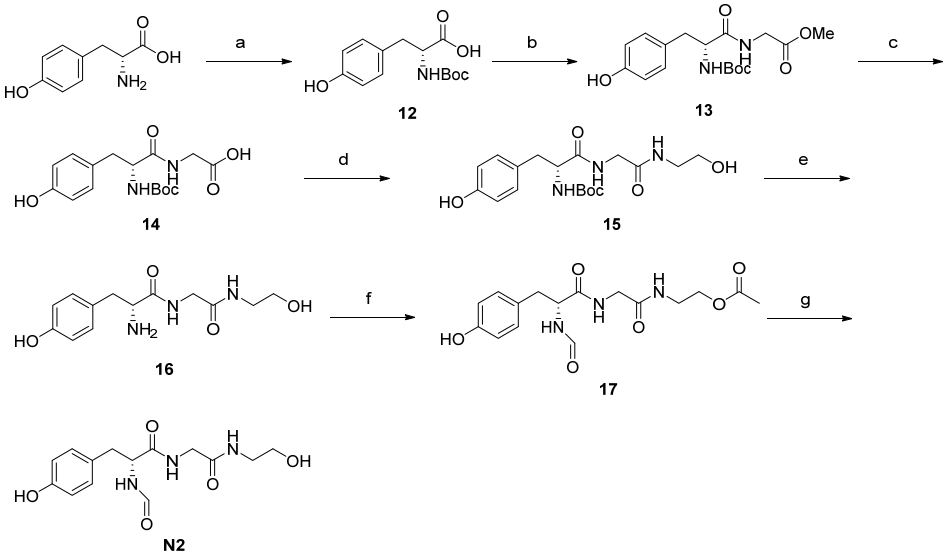

[a]Reagents and conditions: a) Fmoc-Cl, $H_2O$, dioxane, 25 °C, 12 h, 81%; b) $SOCl_2$, DCM, 25 °C, 12 h, 71%; c) *trans*-4-(aminomethyl)cyclohexanecarboxylic acid, 10% NaOH, $H_2O$, 25 °C, 12 h, 77%; d) $HCO_2H$, $Ac_2O$, 25 °C, 12 h, 62%.

### 4.3.4 Synthesis of compound N1 and derivatives thereof

The stereocenter present in compound **N1** was introduced by the commercially available D-tyrosine, which was protected to afford Boc-D-tyrosine (**12**) and coupled to glycine methyl ester in the presence of NMM and CDMT obtaining dipeptide Boc-D-Tyr-D-Gly-OMe (**13**, Scheme 4). Dipeptide **13** was hydrolyzed under basic conditions to obtain Boc-D-Tyr-Gly (**14**), followed by a coupling reaction with ethanolamine in the presence of DCC as a coupling agent, which provided Boc-D-Tyr-Gly-aminoethanol (**15**). Dipeptide **15** was deprotected to give D-Tyr-Gly-aminoethanol (**16**) and then reacted with acetic acid and acetic anhydride affording the desired formamide **17**, whose acetyl group was removed in the presence of $K_2CO_3$ in methanol delivering alcohol **N2**.

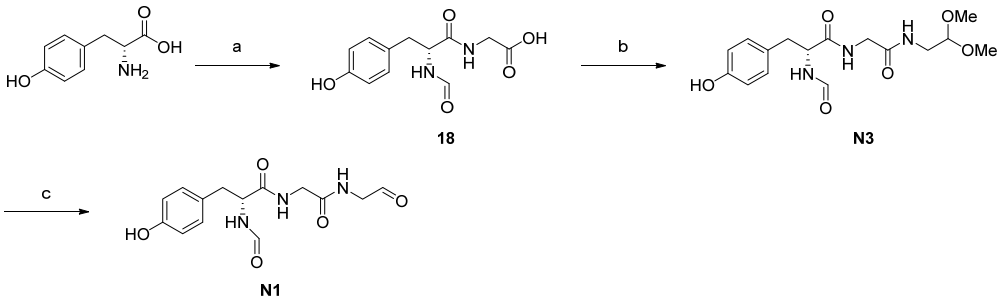
Unfortunately, alcohol **N2** could not be successfully oxidized into the final product **N1**. Thus, a more convergent synthesis was designed (Scheme 5) that comprised the presence of the aldehyde motif masked as an acetal protecting group. However, the synthesized aldehyde **N1** decomposed upon storage at -20 °C, and thus, no biological evaluation was performed.

**Scheme 4**[a]




[a]Reagents and conditions: a) $(Boc)_2O$, $Et_3N$, $H_2O$, dioxane, 25 °C, 12 h, 97%; b) Glycine methyl ester hydrochloride, NMM, CDMT, EtOAc, 25 °C, 3 h, 78%; c) 2M NaOH, MeOH, 25 °C, 2 h, 98 %; d) Ethanolamine, DCC, HOBt, DMF, 25 °C, 12 h, 69%; e) 33% HCl, EtOH, $H_2O$, 25 °C, 12 h, 98%; f) $HCO_2H$, $Ac_2O$, 25 °C, 12 h, 41 % ; g) $K_2CO_3$, MeOH, quantitative.

**Scheme 5**[a]



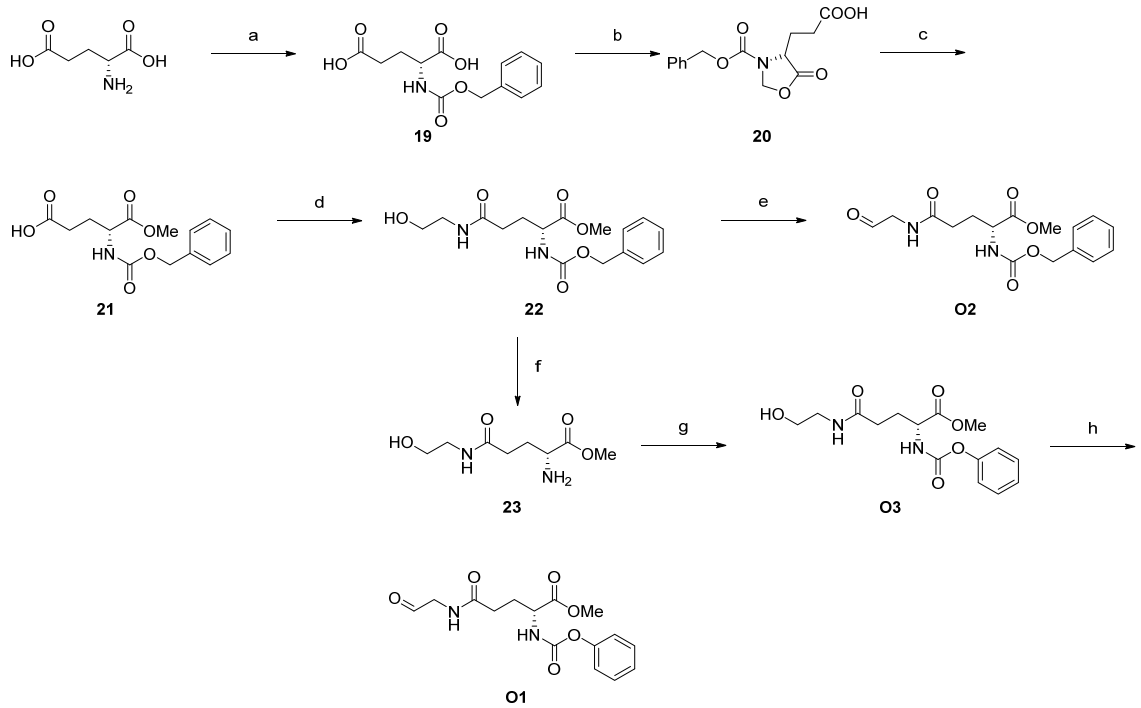

[a]Reagents and conditions: a) (i) $Ac_2O$, $HCO_2H$, 2 h, 86%, (ii) Glycine methyl ester, DCC, HOBt, $Et_3N$, DMF, 25 °C, 12 h, 71%, (iii) 2M NaOH, MeOH, rt, 2 h, 100%; b) Aminoacetaldehyde dimethyl acetal, DCC, HOBt, DMF, 25 °C, 12 h, 49%; c) 15% HCl, $H_2O$, 25 °C, 3h, 35%.

The corresponding enantiomers of **N1**-**N3**, referred to as **N4**-**N6**, were prepared in an analogous manner to that described in Schemes 4 and 5 starting from the commercially available L-tyrosine.

### 4.3.5 Synthesis of compound O1 and derivatives thereof

The synthesis of compound **O1** started with the protection of D-glutamic acid with benzyl chloroformate to give *N*-Cbz-glutamic acid **19**. The protected compound was reacted with paraformaldehyde in the presence of *p*-toluensulfonic acid to obtain **20**, which was converted into 1-methyl-*N*-Cbz-glutamic ester (**21**) by reacting it with NaOMe. The monomethyl ester **21** was then coupled to ethanolamine in the presence of DCC and HOBt to give *N*-ethanol-1-methyl-*N*-Cbz-glutamic ester (**22**). The primary alcohol in compound **22** was oxidized to aldehyde **O2**. In order to achieve the desired compound **O1**, the Cbz protected compound **22** was deprotected and reacted with phenyl choloroformate obtaining alcohol **O3**, which was subsequently oxidized in the presence of DMP affording the final aldehyde **O1**.

**Scheme 6**[a]




[a]Reagents and conditions: a) Benzyl chloroformate, 4M NaOH, 60 °C, 15 min, 38% ; b) Paraformaldehyde, *p*-TsOH, toluene, reflux, 3 h, 91%; c) 1M NaOMe, MeOH, 0 °C, 30 min, 80%; d) Ethanolamine, DCC, HOBt, DMF, 25 °C, 12 h, 41%; e) DMP, DCM, 25 °C, 1 h, 41%; f) 10% Pd/C, $H_2$, MeOH, 25 °C, 81 %; g) Phenyl chloroformate, $Et_3N$, dioaxane, 25 °C, 12 h, 18%; h) DMP, DCM, 0 °C, 1 h, 46%.

The corresponding enantiomers of **O1**-**O3**, referred to as **O4**-**O6**, were prepared in an analogous manner to that shown in Scheme 6 starting from the commercially available L-glutamic acid.

## 4.4 Biological evaluation

An overview of the tested molecules is given in Figure 22.

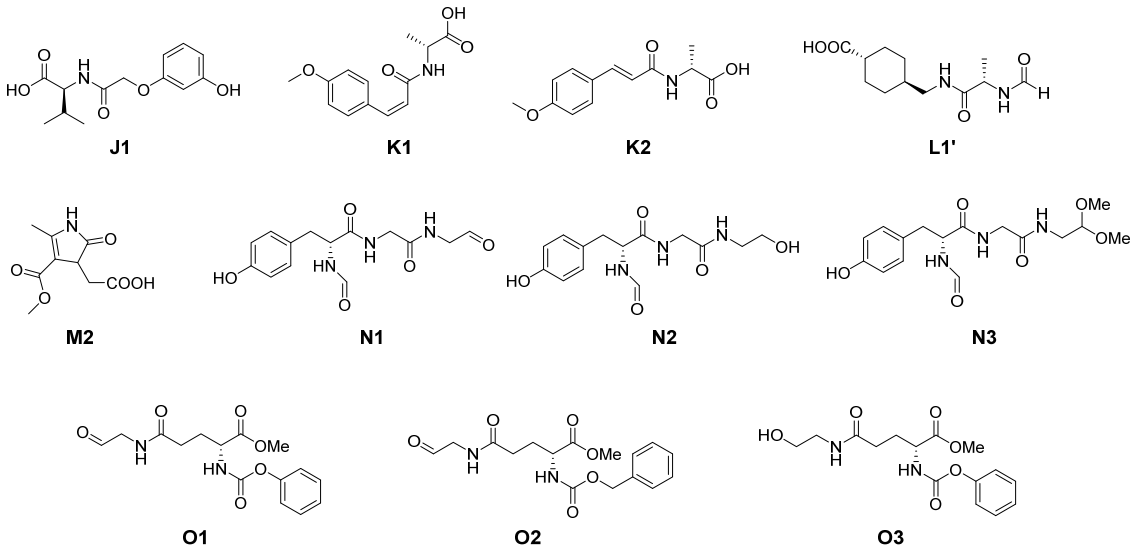



**Figure 22.** Overview of the tested compounds.

First, cell viability in the presence of these compounds was conducted in four different cell lines: 3T3 NIH mouse fibroblasts, HT-29 colon adenocarcinoma, MDA-MD-231 breast carcinoma and A549 lung adenocarcinoma cells. Cell viability was studied by measuring the ability of the cells to process resazurin, which becomes fluorescent under mitochondrial reduction and its emission directly correlates with the metabolic viability of the cells. No toxicity was observed for any of the tested molecules (Figure 22) upon incubation of the inhibitors for 4, 24 or 48 hours at concentrations as high as 100 μM, confirming their safe toxicity profile.

Next, the effects of the molecules on the actin cytoskeleton were studied. For this purpose, a well-established assay involves the incubation of NIH/3T3 fibroblasts with the drug of choice. Morphological changes can easily be monitored by fluorescence marked phalloidin staining of actin that provides a straightforward read-out of the impact of such molecules on the actin network by fluorescence microscopy. The compounds showing an impact on the actin cytoskeleton and their effects are given in Figure 23.

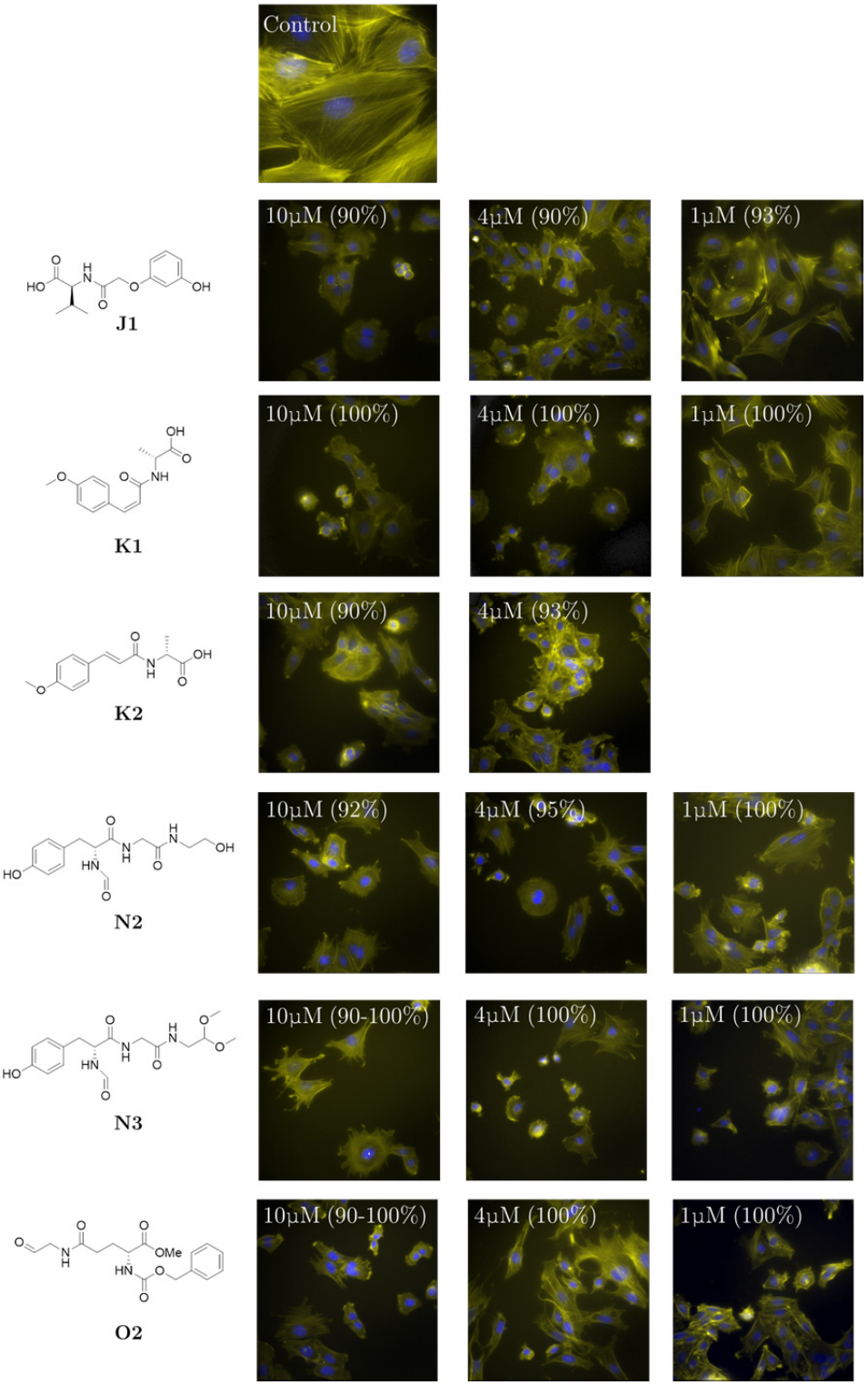



**Figure 23.** Effect of compounds **K1**, **K2**, **N2**, **N3** and **O2** on mouse fibroblasts after 4 hours of incubation. Fluorescence micrographs (40x) of the actin cytoskeleton of NIH/3T3 cells: actin cytoeskeleton is stained with FITC-phalloidin (yellow) and nuclei with 2-(4-amidinophenyl)-6-indolecarbamidine hydrochloride (DAPI) (blue). The value in (%) indicates the percentage of viable cells.

At low micromolar concentrations the compounds shown in Figure 23 elicited considerable morphological changes in cells starting only 4 hours after compound application. The inhibitors caused rounding of NIH/3T3 cells and loss of cell contacts. Moreover, well spread cells lost their smooth contour and a diminution of the microfilament bundles (stress fibers), together with cell shrinkage, was also observed. At higher concentrations of the drug (4 and 10 μM), depletion of F-actin from the central region and formation of actin bundles at the cell margins was observed (Figure 23). Interestingly, many of the cells became binuclear, indicating that the molecules did not affect mitosis but most likely inhibit cytokinesis. Thus, the molecules shown in Figure 23 were able to modify the actin cytoskeleton of fibroblast cells at non-toxic concentrations, provoking morphological changes that resemble that of actin binding natural products.

To confirm the binding of the design inhibitors towards actin, an *in vitro* actin polymerization fluorescence assay was conducted with pyrene labelled actin monomers. The time course of polymerization can be monitored by the fluorescence increase of pyrene during the copolymerization of actin with fluorescence labelled (pyrene) actin.[84] Generally, a lag time is observed (nucleation step) followed by an exponential phase (elongation) and a plateau when the steady state is reached. At a concentration of 10 μM, the tested compounds were able to slightly inhibit the polymerization of actin by decreasing the extent of polymerization. Unfortunately, their inhibition values lied far away to those of the natural product latrunculin (Figure 24).

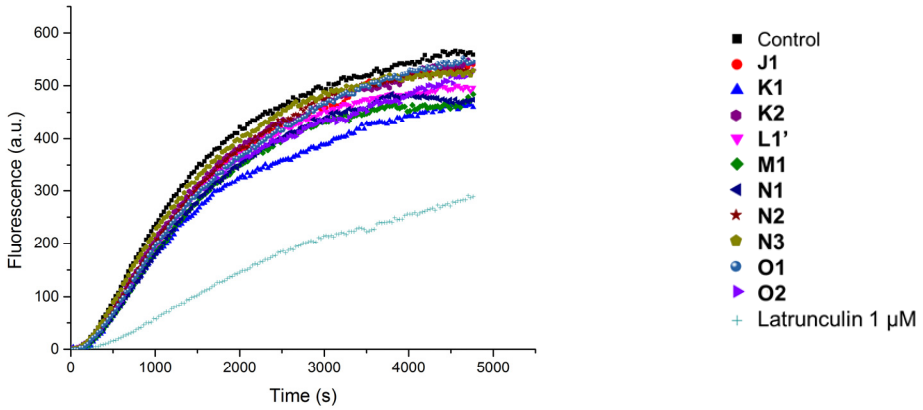



**Figure 24.** Effect of compounds **J**-**O** on actin polymerization *in vitro*. Actin/pyrenyl-labeled actin (2.5 μM) was incubated with the corresponding compounds (10 μM) or DMSO as control (shown in black squares). Polymerization was started by the addition of inducing salts (50 mM KCl, 2 mM $MgCl_2$, 1 mM ATP) and drugs dissolved in DMSO. Latrunculin is given as a reference compound at 1 μM concentration.

We succeeded in recording a dose response curve for compound **O1** (Figure 25), confirming its over 100 fold weaker inhibition in comparison to latrunculin. However, it is

worth underlying the ease of synthesis of the designed actin inhibitors in respect to the complex natural product latrunculin (26 steps with an overall yield of 0.21%),[85] making our compounds readily accessible in high quantities.

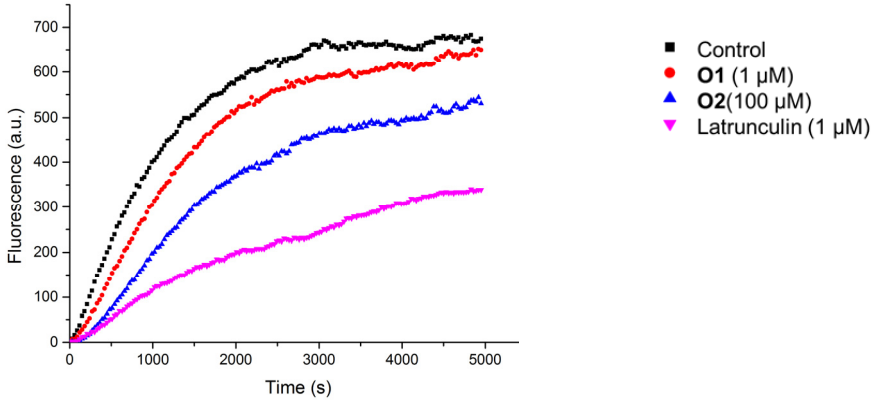



**Figure 25.** Concentration dependent effect of compound **O1** on actin polymerization *in vitro*. Actin/pyrenyl-labeled actin (2.5 µM) was incubated with **O1** at 1 and 100 µM or DMSO as control (shown in black squares). Polymerization was started by the addition of inducing salts (50 mM KCl, 2 mM $MgCl_2$, 1 mM ATP) and drugs dissolved in DMSO. Latrunculin is given as a reference compound at 1 µM concentration.

## 4.5 Conclusions and outlook

We have designed and validated a new computational approach to discover novel actin leads targeting the ATP binding site of actin. Several fragments were selected from an in house constructed library and ranked by means of consensus scoring. The best ranking compounds of five different consensus scorings were visually analyzed and fifteen fragments were selected and subjected to MD simulations, where the binding behavior was studied. From those fifteen fragments, five scaffolds were chemically synthesized.

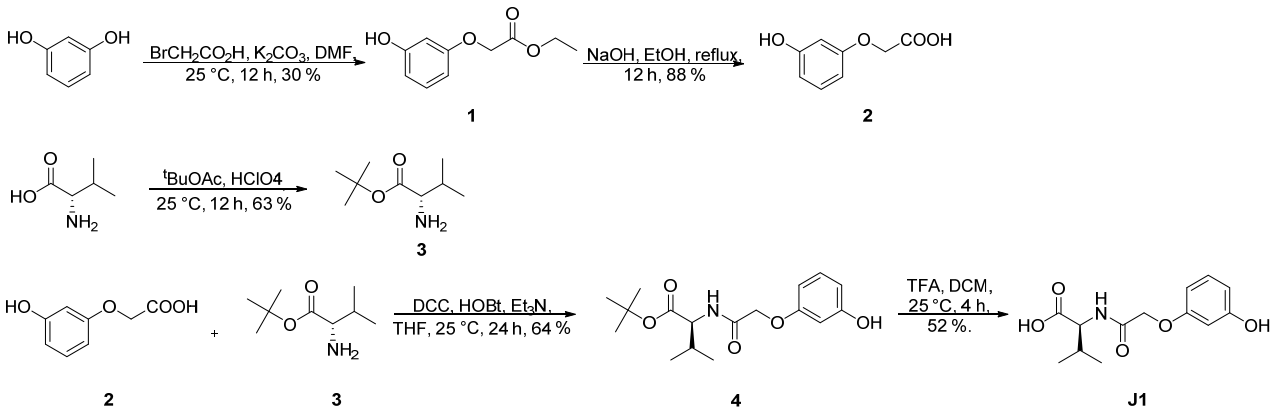
Cytotoxicity studies on NIH/3T3 fibroblasts and several human cancer cells by the colorimetric rezazurin assay revealed low toxicity values for the tested molecules. The actin cytoskeleton of the NIH/3T3 cells was then stained with TRITC-Phalloidin to visualize the morphological changes in the actin cytoskeleton. The molecules were able to alter the morphology of the actin cytoskeleton at low micromolar concentrations and after a short incubation time (4 hours) by reducing the microfilament bundles (stress fibers), inducing cell shrinkage and actin aggregation. To confirm the binding of the design inhibitors towards actin, an *in vitro* actin polymerization fluorescence assay was conducted with pyrene labelled actin monomers, obtaining a weak inhibition of actin polymerization.

As a result, the developed small organic molecules constitute valuable tools for the study of actin dynamics and are promising starting hits for the development of more potent actin inhibitors. The next steps of the project include the design and synthesis of more potent derivatives of the mentioned scaffolds, as well as the implementation of computer based *de novo* design aiming to develop novel actin binding molecules.

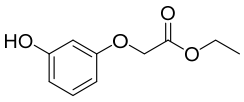
## 4.6 Experimental section

### 4.6.1 Chemistry. General methods

All reactions, unless otherwise stated, were carried out under a nitrogen atmosphere using standard Schlenk-techniques. All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 F254. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV 300, AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in ppm. The spectra are calibrated to the residual 1H and 13C signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Buchi melting point B-540 instrument. High-resolution electrospray ionization mass spectrometry was performed on a Finnigan MAT 900 (Thermo Finnigan, San Jose, CA, USA) double-focusing magnetic sector mass spectrometer. Ten spectra were acquired. A mass accuracy ≤2 ppm was obtained in the peak matching acquisition mode by using a solution containing 2 µL PEG200, 2 µL PPG450, and 1.5 mg NaOAc (all obtained from Sigma-Aldrich, Buchs, Switzerland) dissolved in 100 mL MeOH (HPLC Supra grade, Scharlau, E-Barcelona) as internal standard. The purity of all tested compounds was determined by HPLC on a Waters Acquity UPLC (Waters, Milford, MA) Top spectrometer using an Acquity BEH C18 HPLC column (1.7 $\mu$ m, 1× 50 mm, Waters) with a mixture of $H_2O$ + 0.1% HCOOH (A) and $CH_3CN$ + 0.1% HCOOH (B) solvent (0.1 mL flow rate, linear gradient from 5% to 98% B within 4 min followed by flushing with 98% B for 1 min). Unless otherwise stated, all compounds showed ≥ 95 % purity.

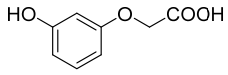



**Ethyl 2-(3-hydroxyphenoxy)acetate (1)**[86]



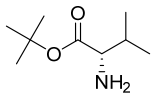
To a mixture of resorcinol (0.50 g, 4.54 mmol) and $K_2CO_3$ (0.63 g, 4.54 mmol) in DMF (8.0 mL), ethyl bromoacetate (0.51 mL, 4.54 mmol) was added dropwise at 0 °C. The mixture was stirred at 25 °C for 12 hours, concentrated and redissolved in EtOAc. The insoluble solid was filtered off and the solution was extracted with 10 % citric acid, 5 % sodium bicarbonate and brine. The organic phase was dried over $MgSO_4$ and evaporated under reduced pressure. The product was purified by flash column chromatography on silica gel (hexane/EtOAc, 7:1) to provide a colorless oil (0.27 g, 30 % yield). $^1$H NMR (500 MHz, $CDCl_3$): δ= 6.97 (t, $J$ = 8.2 Hz, 1H), 6.72 (br, 1H), 6.40 (ddd, $J$ = 8.1, 2.2, 0.7 Hz, 1H), 6.36 (t, $J$ = 2.2 Hz, 1H), 6.32-6.35 (m, 1H), 4.48 (s, 2H), 4.17 (q, $J$ = 7.1 Hz, 2H), 1.20 (t, $J$ = 7.1 Hz, 3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 169.6, 158.7, 157.1, 130.1, 109.2, 106.1, 102.6, 65.2, 61.6, 14.0; IR (neat): $\tilde{\upsilon}$ = 3422, 2988, 2908, 1749, 1603, 1594, 1514, 1464, 1435, 1223, 1161, 1092, 1021, 951, 864, 762, 686, 599 cm$^{-1}$ ; HRMS (ESI), *m/z*: calcd for $C_{10}H_{12}NaO_4$, 219.0628; found, 219.0628 $[M + Na]^+$.

**2-(3-Hydroxyphenoxy)acetic acid (2)**



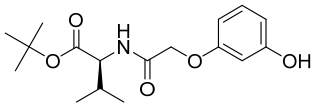
Ethyl 2-(3-hydroxyphenoxy)acetate (0.77 g, 3.93 mmol) was dissolved in methanol (7.0 mL) and a 2M NaOH aqueous solution (4.0 mL) was added at 0 °C. The mixture was stirred at 25 °C until the dissapearance of the starting material. The reaction mixture was neutralized by the addition of 1M HCl solution and extracted with DCM. The pH was adjusted to 2 by 1M HCl addition and extracted with EtOAc. The organic phases were dried over $MgSO_4$ and evaporated to afford the product as a white solid (88 % yield). $^1$H NMR (500 MHz, DMSO-$d_6$): δ= 12.94 (br, 1H), 9.41 (s, 1H), 7.04 (t, $J$ = 8.1 Hz, 1H), 6.36 (ddd, $J$ = 8.1, 2.2, 0.8 Hz, 1H), 6.32 (ddd, $J$ = 8.2 , 2.5, 0.8 Hz, 1H), 6.28 (t, $J$ = 2.3 Hz, 1H), 4.57 (s, 2H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 170.2, 158.8, 158.4, 129.8, 108.2, 104.9, 101.6, 64.3; IR (neat): $\tilde{\upsilon}$ = 3479, 2970, 1714, 1589, 1499, 1455, 1431, 1272, 1246, 1175, 1146, 1088, 966, 818, 687 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_8H_8NaO_4$, 191.0315; found, 191.0314 $[M + Na]^+$.

**L-Valine *tert*-butyl ester (3)**[87]



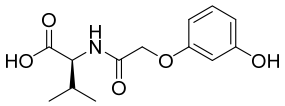
To a solution of L-valine (1.00 g, 17.07 mmol) in $^{t}$BuAc (20.0 mL), $HClO_4$ (2.51 g, 25.00 mmol) were added dropwise at 0 °C and the mixture was stirred at 25 °C for 12 hours. The reaction mixture was washed with water and 1M HCl solution. The aqueous solution was adjusted to pH 9 by adding a 10 % $K_2CO_3$ solution and it was extracted with DCM. The combined organic extracts were dried over $MgSO_4$. Purification by column chromatography on silica gel (hexanes/EtOAc, 3:1) afforded the desired compound as a colorless oil (63 %). $^1$H NMR (400 MHz, $CDCl_3$): δ= 3.13 (d, $J$ = 4.8 Hz, 1H), 1.92-2.00 (m, 1H), 1.43 (s, 9H), 1.36 (s, 2H), 0.94 (dd, $J$ = 6.7, 5.2 Hz, 3H), 0.87 (dd, $J$ = 6.7, 5.2 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ= 174.8, 80.7, 60.3, 32.1, 28.0, 19.2, 17.0; IR (film): $\tilde{\upsilon}$ = 3391, 3316, 2964, 2933, 2875, 1725, 1460, 1391, 1367, 1251, 1147, 973, 850 cm$^{-1}$; MS (ESI): *m/z*: calcd for $C_9H_{19}NNaO_4$: 196.1, found: 196.0 $[M + Na]^+$.

***N*-2-(3-Hydroxyphenoxy)acetoyl-L-valine *tert*-butyl ester (4)**



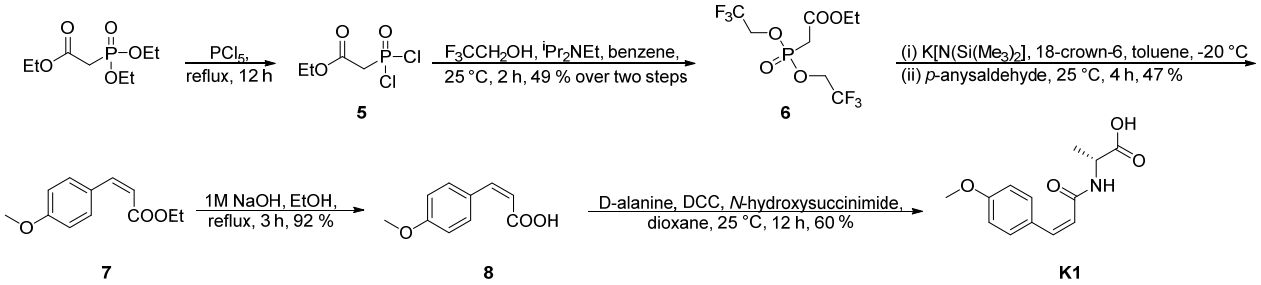
To a solution of 2-(3-hydroxyphenoxy)acetic acid (**2**) (0.17 g, 1.00 mmol) in THF (10.0 mL) was added L-valine tert-butyl ester (**3**) (0.17 g, 1.00 mmol), HOBt (0.14 g, 1.00 mmol) and $Et_3N$ (138 µL, 1.00 mmol) in 20.0 mL THF. After stirring the solution for 15 min, DCC (0.25 g, 1.20 mmol) was added in 4.0 mL THF. The reaction mixture was stirred at 25 °C for 24 hours. The dicyclohexyl urea formed was filtered off and the residue was purified by flash column chromatography (3:1 hexane/EtOAc) to afford the desired product as an orange oil (64 % yield). $^1$H NMR (500 MHz, $CDCl_3$): δ= 7.10-7.15 (m, 2H), 6.53 (dd, $J$ = 7.8, 1.2 Hz, 1H), 6.46 (dd, $J$ = 10.5, 2.2 Hz, 2H), 4.48-4.53 (m, 3H), 2.19 (qd, $J$ = 13.5, 6.8 Hz, 1H), 1.46 (s, 9H), 0.93 (d, $J$ = 6.9 Hz, 3H), 0.89 (d, $J$ = 6.9 Hz, 3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 170.6, 168.7, 158.2, 157.9, 130.3, 109.4, 106.4, 102.3, 82.5, 67.0, 57.2, 31.5, 28.0, 18.8, 17.5; IR (neat): $\tilde{\upsilon}$ = 3338, 2969, 2933, 1729, 1665, 1597, 1530, 1489, 1369, 1285, 1145, 911, 842, 731, 686 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{17}H_{25}NNaO_5$, 346.1625; found, 346.1622 $[M + Na]^+$; $[\alpha]_D^{26}$ + 16.1 (c 1.00, DCM).

***N*-2-(3-Hydroxyphenoxy)acetoyl-L-valine (J1)**

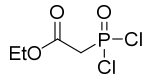


To a solution of *N*-2-(3-hydroxyphenoxy)acetoyl-L-valine tert-butyl ester (**4**, 119 mg, 0.37 mmol) in DCM (2.0 mL) was added TFA in excess (0.73 mL, 9.60 mmol) at 0 °C. The reaction was stirred at 25 °C for 4h. It was concentrated and the residue was dissolved in water, which was extracted with DCM. The pH of the aqueous phase was adjusted to 2 by the addition of 1M HCl and it was extracted with EtOAc. The EtOAc extracts were dried over $MgSO_4$ and concentrated in vaccuo to afford the product in a 52 % yield as a yellow solid. $^1$H

NMR (500 MHz, MeOD): δ= 7.09 (t, $J$ = 8.3 Hz, 1H), 6.46 (ddd, $J$ = 8.2, 2.3, 0.9 Hz, 1H), 6.42-6.45 (m, 2H), 4.54 (s, 2H), 4.38 (d, $J$ = 5.0 Hz, 1H), 2.18-2.24 (m, 1H), 0.94 (d, $J$ = 6.9 Hz, 3H), 0.90 (d, $J$ = 6.9 Hz, 3H); $^{13}$C NMR (125 MHz, MeOD): δ= 171.0, 160.3, 160.0, 131.2, 110.2, 106.7, 103.4, 68.1, 32.3, 19.7, 18.1, one carbon is missing due to overlapping; IR (neat): $\tilde{\upsilon}$ = 3326, 2965, 1715, 1658, 1597, 1536, 1284, 1173, 1147, 1059, 837, 762, 684 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{17}NNaO_5$, 290.0999; found, 290.0997 $[M + Na]^+$; $[\alpha]_D^{26}$ + 3.5 (c 0.67, MeOH).

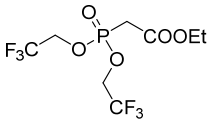



**Ethyl P,P-bischloro phosphonoacetate (5)**[88]



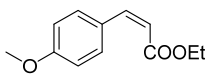
Triethyl phosphonoacetate (2.00 g, 8.92 mmol) was cooled to 0 °C and phosphorus (V) pentachloride (4.64 g, 22.30 mmol) was added. The reaction mixture was stirred for 1 hour at 25 °C and heated to 75 °C for 12 hours. The excess of $PCl_5$ and $POCl_3$ were removed under vacuum and the bischloro phosphonoacetate (yellow oil, 84 % yield) was used without further purification in the next step.

**Ethyl P,P-bis(2,2,2-trifluoroethyl) phosphonoacetate (6)**[88]



Dichloride **5** (0.50 g, 2.44 mmol) was dissolved in benzene (3.0 mL) and a solution of trifluoroethanol (0.35 mL, 4.87 mmol) and diisopropyl amine (0.85 mL, 4.87 mmol) in benzene (4.30 mL) was added at 0 °C. The solution was stirred at 25 °C for 2 hours. Evaporation of the solvent and column chromatography on silica gel (hexane/EtOAc, 3:1) afforded the desired compound as a colorless oil (49 % yield). $^1$H NMR (400 MHz, $CDCl_3$): δ= 4.43 (quint, $J$ = 8.3 Hz, 4H), 4.19 (q, $J$ = 7.1 Hz, 2H), 3.12 (d, $J$ = 21.2 Hz, 2H), 1.26 (t, $J$ = 7.2 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ= 164.6 (d, $J$ = 4.3 Hz), 122.4 (qd, $J$ = 277.3, 8.4 Hz), 62.5 (qd, $J$ = 38.2, 5.5 Hz), 62.2, 33.9 (d, $J$ = 143.9 Hz), 13.8; IR (film): 2985, 2927, 1737, 1420, 1287, 1260, 1163, 1096, 1065, 1028, 961, 897, 843, 652, 553, 471 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_8H_{11}F_6NaO_5P$, 355.0; found, 355.0 $[M + Na]^+$.

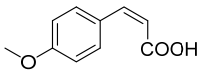
***Cis*-4-Methoxycinnamate (7)**



A solution of the phosphonoacetate **6** (0.80 g, 2.41 mmol) and 15-crown-5 (0.43 g, 1.61 mmol) in dry THF (12.0 mL) was cooled to -78 °C under nitrogen. 1 hour after adding a 0.5 M solution of KHMDS in toluene (4.8 mL), freshly distilled *p*-anysaldehyde (0.33 g, 2.41 mmol) was added by slow dropwise. The

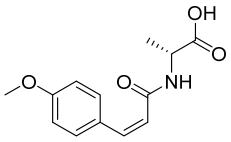
reaction was stirred at -20 °C for 1 hour and at 25 °C for 3 hours. The reaction was quenched by addition of saturated ammonium chloride. The aqueous layer was then extracted with $Et_2O$ three times, the combined organic phases were dried over $MgSO_4$ and the diastereoisomers were separated by flash column chromatography (toluene, 47 % yield). $^1$H NMR (400 MHz, $CDCl_3$): δ= 7.69 (d, $J$ = 8.6 Hz, 2H), 6.87 (d, $J$ = 8.9 Hz, 2H), 6.83 (d, $J$ = 12.8 Hz, 1H), 5.82 (d, $J$ = 12.7 Hz, 1H), 4.19 (q, $J$ = 7.1 Hz, 2H), 3.81 (s, 3H), 1.28 (t, $J$ = 7.1 Hz, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ= 166.3, 160.3, 143.0, 132.1, 127.3, 117.1, 113.3, 60.0, 55.1, 14.1; IR (film): 2979, 2904, 2839, 1713, 1624, 1601, 1509, 1462, 1441, 1306, 1255, 1190, 1028, 947, 846, 519 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{14}NaO_3$, 229.0835; found, 229.0835 $[M + Na]^+$.

### *Cis*-4-Methoxycinnamic acid (8)



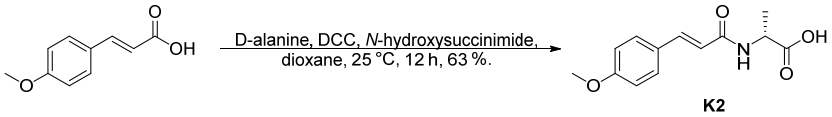

Cis-4-methoxycinnamate **7** (0.20 g, 0.97 mmol) was dissolved in 1M aqueous NaOH (4.0 mL) solution and ethanol (8.0 mL). The reaction mixture was stirred for 2 hours. The pH was neutralized and the solvent evaporated. The ressidue was redissolved in water. The pH was acidified to 2 by addition of 1M HCl and it was then extracted with EtOAc. The combined organic phases were dried over $MgSO_4$ and the acid **8** was obtained in 92 % yield. $^1$H NMR (400 MHz, $CDCl_3$): δ= 7.63 (d, $J$ = 8.7 Hz, 2H), 6.89 (d, $J$ = 12.8 Hz, 1H), 6.81 (d, $J$ = 8.7 Hz, 2H), 5.77 (d, $J$ = 12.8 Hz, 1H), 3.76 (s, 3H); $^{13}$C NMR (100 MHz, $CDCl_3$): δ= 171.5, 160.7, 145.9, 132.6, 127.0, 115.9, 113.5, 55.3; IR (film): 2956, 2936, 2838, 2574, 1684, 1618, 1597, 1509, 1453, 1442, 1306, 1250, 1223, 1168, 1117, 1028, 919, 841, 774, 705, 590, 515 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{10}H_{10}NaO_3$, 201.0522; found, 201.0520 $[M + Na]^+$.

### *N*-(*cis-p*-Methoxycinnamoyl)-D-alanine (K1)

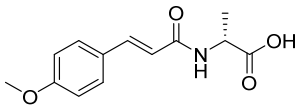



DCC (66.0 mg, 0.320 mmol) was added to a solution of cis-4-methoxycinnamic acid (52.0 mg, 0.294 mmol) and N-hydroxysuccinimide (37 mg, 1.24 mmol) in dioxane (0.5 mL) at 0 °C. The reaction mixture was warmed up to 25 °C and stirred for 5 hours. The insoluble solid was filtered off and washed with cold dioxane. The filtrate was added to a solution of D-alanine (29.0 mg, 0.326 mmol) and sodium bicarbonate (25.0 mg) in water (0.4 mL) and it was stirred at 25 °C for 24 hours. The solvent was evaporated and the residue redissolved in water. The solid was filtered off and the aqueous phase was extracted with EtOAc. The organic extracts were discarded, the aqueous phase was adjusted to pH2 by addition of concentrated HCl and it was extracted with EtOAc. The combined organic extracts were dried over $MgSO_4$. The desired compound was purified by flash column chromatography ($Et_2O$), obtaining a white solid in 60% yield. $^1$H NMR (500 MHz, $CDCl_3$): δ= 7.49 (d, $J$ = 8.8 Hz, 2H), 6.87 (d, $J$ = 8.8 Hz, 2H), 6.79 (d, $J$ = 12.5 Hz, 1H), 6.07 (d, $J$ = 6.6 Hz, 1H), 5.86 (d, $J$ = 12.5 Hz, 1H), 4.60 (quint, $J$ = 7.1 Hz, 1H), 3.81 (s, 3H), 1.39 (d, $J$ = 7.2 Hz,

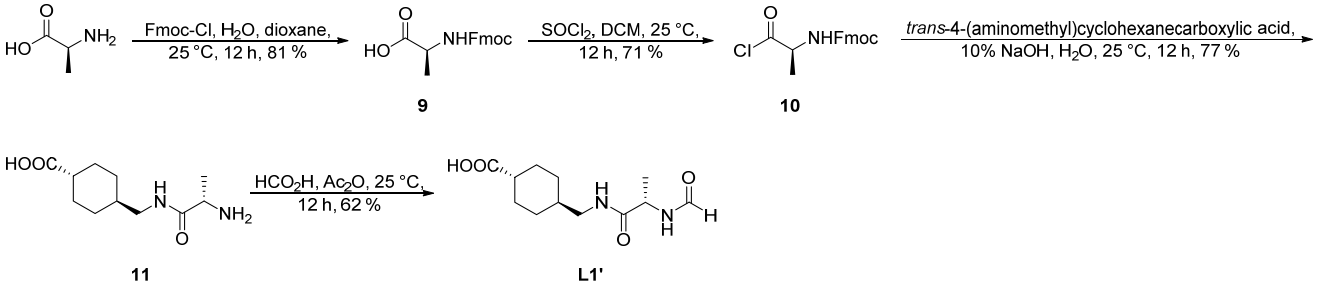
3H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 175.4, 167.5, 160.2, 138.7, 131.1, 127.0, 120.9, 113.8, 55.3, 48.3, 17.4; IR (film): 3330, 2936, 1734, 1646, 1604, 1511, 1456, 1257, 1175, 1030, 838, 750 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{15}NNaO_4$, 272.0893; found, 272.0889 $[M + Na]^+$; $[\alpha]_D^{26}$ + 49.6 (c 1.14, DCM).



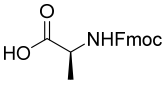

### *N*-(*Trans-p*-methoxycinnamoyl)-D-alanine (K2)[89]



Following an identical procedure to the one described for compound **K1**, starting from trans-4-methoxycinnamic acid (0.20 g, 1.12 mmol), **K2** was obtained as a white solid (63 % yield). $^1$H NMR (400 MHz, DMSO-$d_6$): δ= 12.57 (br, 1H), 8.28 (d, $J$ = 7.4 Hz, 1H), 7.51 (d, $J$ = 8.6 Hz, 2H), 7.38 (d, $J$ = 15.8 Hz, 1H), 6.98 (d, $J$ = 8.5 Hz, 2H), 6.55 (d, $J$ = 15.8 Hz, 1H), 4.33 (quint, $J$ = 7.3 Hz, 1H), 3.79 (s, 3H), 1.31 (d, $J$ = 7.3 Hz, 3H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ= 174.1, 164.8, 160.3, 138.7, 129.0, 127.3, 119.2, 114.3, 55.2, 47.5, 17.2; IR (neat): 3374, 2985, 2945, 1716, 1650, 1595, 1523, 1509, 1459, 1389, 1320, 1249, 1207, 1176, 1154, 1110, 1040, 994, 895, 827, 648, 577, 538, 513, 422 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{13}H_{14}NO_4$, 248.1; found, 247.9 $[M - H]^-$; $[\alpha]_D^{26}$ – 27.1 (c 0.95, MeOH).



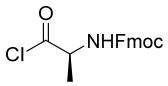

### *N*-Fmoc-L-alanine (9)[90]



L-alanine (1.00 g, 11.22 mmol) was dissolved in water (44.0 mL) and $Na_2CO_3$ (3.0 g) was added. The solution was cooled to 0 °C and dioxane (26.0 mL) was added. Fmoc-Cl (3.19 g, 12.35 mmol), dissolved in 26.0 mL of dioxane, was added dropwise. The reaction mixture was stirred at 0 °C for two hours and at 25 °C for 12 hours. The solvent was evaporated under reduced preassure and the obtained solid was redissolved in water. The aqueous phase was extracted with diethyl ether. Then, the aqueous phase was adicified to pH 2 by the addition of 1M HCl solution and it was extracted with ethyl acetate three times. The combined EtOAc phases were dried over $MgSO_4$ and concentrated under vacuum. The obtained product was recrystallized from DCM/hexanes to afford a white solid (81 % yield). $^1$H NMR (400 MHz, DMSO-$d_6$): δ= 12.52 (br, 1H), 7.89 (d, $J$ = 7.5 Hz, 2H), 7.72 (dd, $J$ = 7.3, 3.8 Hz, 2H), 7.65 (d, $J$ = 7.5 Hz, 1H), 7.42 (t, $J$ = 7.5 Hz, 2H), 7.33 (t, $J$ = 7.4 Hz, 2H), 4.20-4.29 (m, 3H), 4.00 (quint, $J$ = 7.2 Hz, 1H), 1.28 (d, $J$ = 7.4 Hz, 3H); $^{13}$C NMR (100 MHz, DMSO-$d_6$): δ= 174.2,

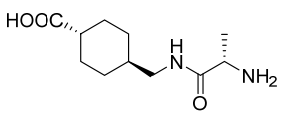
155.7, 143.7, 143.7, 140.6, 127.5, 127.0, 125.1, 120.0, 65.5, 49.1, 46.5, 17.0; IR (neat): $\tilde{\upsilon}$ = 3513, 3414, 3340, 2985, 1717, 1702, 1678, 1528, 1449, 1307, 1246, 1082, 1032, 765, 743 $cm^{-1}$; MS (ESI): *m/z*: calcd for $C_{18}H_{17}O_4$: 310.1, found: 310.0 $[M - H]^-$.

**_N_-Fmoc-L-alanine acid chloride (10)**[91]



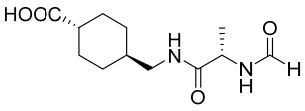

To a solution of N-Fmoc-L-alanine (**9**, 1.00 g, 3.24 mmol) in DCM (13.0 mL) was added $SOCl_2$ (2.3 mL, 32.40 mmol). The reaction was stirred at 25 °C for 17 hours.The excess of $SOCl_2$ was removed under vacuo and the obtained solid was used in the next step without further purification. (71 % yield).

***Trans*-4-(aminomethyl)cyclohexanecarboxyl-L-alanine (11)**



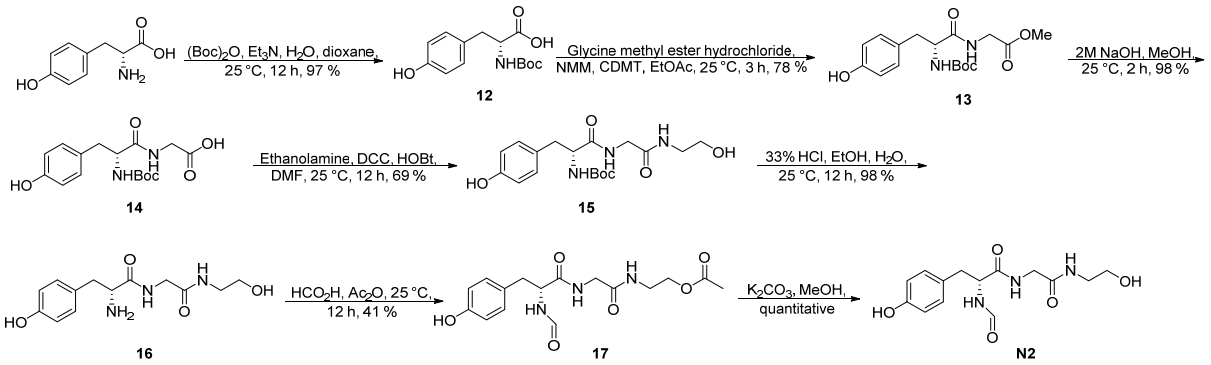

Trans-4-(aminomethyl)cyclohexanecarboxylic acid (172 mg, 1.09 mmol) was dissolved in 10% NaOH aqueous solution (3.0 mL). N-Fmoc-L-alanine chloride (**10**, 359 mg, 1.09 mmol) was added and the solution was stirred for 12 hours. The white precitate was filtered off and the pH of the solution was neutralized by addition of 1M HCl solution. The solvent was evaporated, the obtained solid was dissolved in methanol and the salts were filtered off. The solvent was removed under reduced pressure to afford a white solid (77 % yield). $^1$H NMR (500 MHz, MeOD): δ= 3.58 – 3.79 (m, 1H), 3.06 (dd, *J* = 6.7, 3.3 Hz, 1H), 2.77 (d, *J* = 7.0 Hz, 1H), 2.08 (ddd, *J* = 12.1, 10.9, 6.7 Hz, 1H), 1.94-2.01 (m, 2H), 1.80 – 1.87 (m, 2H), 1.61 (ddq, *J* = 11.1, 7.1, 3.7 Hz, 1H), 1.36-1.48 (m, 6H), 0.94 – 1.08 (m, 2H); $^{13}$C NMR (125 MHz, MeOD): δ= 184.7, 184.4, 173.3, 51.8, 50.7, 48.0, 47.7, 46.7, 46.4, 38.9, 37.1, 31.5, 31.4, 30.9, 30.8, 30.5, 19.1, 17.4, presence of rotamers; IR (neat): $\tilde{\upsilon}$ = 3466, 3281, 3072, 2928, 2850, 1651, 1550, 1448, 1405, 1212, 928, 779 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{11}H_{21}N_2O_3$, 229.1547; found, 229.1544 $[M + H]^+$; $[\alpha]_D^{27}$ - 1.42 (c 1.02, MeOH).

**_N_-Formyl-*trans*-4-(aminomethyl)cyclohexanecarboxyl-L-alanine (L1')**

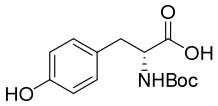



To a solution of **11** (0.10 g, 0.44 mmol) in formic acid (2.3 mL) acetic anhydride (1.3 mL) was added dropwise at 0 °C. The solution was warmed to 25 °C and stirred for 12 hours. The solvent was removed under reduced pressure. The desired compound precipitated in water when the pH was adicified to 2 and it was washed with pentane to afford a white solid (62 % yield). $^1$H NMR (500 MHz, MeOD): δ= 8.08 (s, 1H), 3.94 – 4.50 (m, 1H), 3.09 (dd, *J* = 6.7, 2.1 Hz, 1H), 3.05 (dd, *J* = 6.9, 1.9 Hz, 1H), 2.18 – 2.26 (m, 1H), 1.98 – 2.01 (m, 2H), 1.81 – 1.86 (m, 2H), 1.35 – 1.52 (m, 6H), 0.95 – 1.06 (m, 2H); $^{13}$C NMR (125 MHz, MeOD): δ= 179.9, 179.8, 175.4, 174.6, 171.1, 164.0, 163.5, 163.4, 50.3, 48.0, 46.5, 46.4, 45.1, 44.4, 44.4, 38.6, 38.5, 30.8, 29.8, 18.6, 18.0, 17.9, presence of rotamers; IR (neat): $\tilde{\upsilon}$ = 3270, 3060, 2929, 2864, 1715, 1651, 1540, 1451, 1381,1232, 1193, 1087, 667 $cm^{-1}$; HRMS

(ESI), *m/z*: calcd for $C_{12}H_{20}N_2NaO_4$, 279.1315; found, 279.1313 $[M + Na]^+$; $[\alpha]_D^{26}$ -12.6 (c 1.18, MeOH).

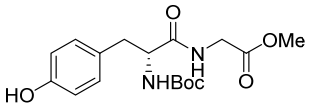



**Boc-D-tyrosine (12)**[92]



To a dispersion of D-tyrosine (2.00 g, 11.04 mmol) in 1/1 dioxane/water (28.0 mL) $Et_3N$ (2.31 mL, 16.56 mmol) was added. The round bottomed flask was cooled to 0 °C and $(Boc)_2O$ (3.61 g, 16.56 mmol) was added. The reaction mixture was stirred at 0 °C for 1 hour and at 25 °C for 12 hours. It was concentrated under reduced pressure and diluted with water. The aqueous layer was first extracted with EtOAc three times and then adjusted to pH 2 with a 1M HCl solution. It was extracted with EtOAc three times and brine. The organic extracts were dried over $MgSO_4$ and evaporation of the solvent gave the desired product as a white solid (97 % yield). $^1$H NMR (500 MHz, DMSO-*d*$_6$): δ= 12.50 (br, 1H), 9.16 (s, 1H), 7.02 (d, $J$ = 8.3 Hz, 2H), 6.97 (d, $J$ = 8.3 Hz, 1H), 6.64 (d, $J$ = 8.4 Hz, 2H), 3.99 (dt, $J$ = 9.8, 4.7 Hz, 1H), 2.88 (dd, $J$ = 13.9, 4.6 Hz, 1H), 2.70 (dd, $J$ = 13.8, 10.1 Hz, 1H), 1.33 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-*d*$_6$): δ= 173.7, 155.8, 155.4, 129.9, 128.0, 114.8, 77.9, 55.5, 35.6, 28.1; IR (neat): $\tilde{\upsilon}$ = 3380, 3162, 1725, 1669, 1612, 16001, 1515, 1402, 1370, 1242, 1162, 1136, 1055, 1032, 822, 783, 653, 595, 532, 490 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{18}NO_5$, 280.1; found, 280.0 $[M - H]^-$; $[\alpha]_D^{23}$ -16.3 (c 1.02, MeOH).

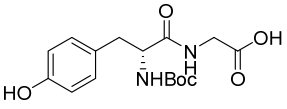
**Boc-D-Tyr-D-Gly-OMe (13)**[93]



N-Methylmorpholine (2.25 g, 22.22 mmol ) was dropwise added to a 0 °C solution of Boc-D-tyrosine (2.50 g, 8.89 mmol), glycine hydrochloride (1.12 g, 8.89 mmol) and CDMT (1.56 g, 8.89 mmol) in EtOAc (25.0 mL). The reaction was stirred at 25 °C for 4 hours. It was then diluted with EtOAc and extracted with water (1x), 1M HCl (2x), saturated $NaHCO_3$ (2x), water (1x) and brine (1x). The organic phase was dried over $MgSO_4$ and concentrated. The residue was purified by flash chromatography on silica gel (2-4 % MeOH in DCM) to afford the desired compound as a white solid (78 % yield). $^1$H NMR (500 MHz, DMSO-*d*$_6$): δ= 9.12 (s, 1H), 8.30 (t, $J$ = 5.8 Hz, 1H), 7.04 (d, $J$ = 8.3 Hz, 2H), 6.80 (d, $J$ = 8.7 Hz, 1H), 6.64 (d, $J$ = 8.4 Hz, 2H), 4.11 (dt, $J$ = 9.9, 4.2 Hz, 1H), 3.86 (ddd, $J$ = 23.1, 17.6, 6.0

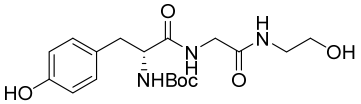
Hz, 2H), 3.63 (s, 3H), 2.87 (dd, $J$ = 14.0, 4.1 Hz, 1H), 2.62 (dd, $J$ = 13.8, 10.3 Hz, 1H), 1.30 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 172.3, 170.1, 155.6, 155.1, 130.0, 128.1, 114.7, 77.8, 55.7, 51.5, 40.5, 36.6, 28.0; IR (neat): ῡ = 3330, 2979, 2937, 1745, 1687, 1660, 1515, 1439, 1365, 1214, 1163, 1019, 826, 541, 445 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{17}H_{24}N_2NaO_6$, 375.1; found, 375.1 $[M + Na]^+$; $[\alpha]_D^{24}$ +0.37 (c 0.98, MeOH).

**Boc-D-Tyr-Gly (14)**[94]



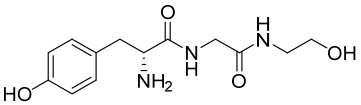

Dipeptide **13** (1.50 g, 4.26 mmol) was dissolved in MeOH (7.5 mL) and cooled to 0 °C. 2M NaOH (4.5 mL) was added dropwise. The reaction mixture was stirred for 3 hours, it was diluted with water and extracted with DCM. The aqueous phase was adicified to pH 2 by 1M HCl addition and extracted it with EtOAc. The organic phase was dried over $MgSO_4$ and the solvent was evaporated under reduced pressure to afford **14** has a white solid (98 % yield). $^1$H NMR (500 MHz, DMSO-$d_6$): δ= 12.54 (br, 1H), 9.11 (s, 1H), 8.15 (t, $J$ = 5.5 Hz, 1H), 7.04 (d, $J$ = 8.3 Hz, 2H), 6.78 (d, $J$ = 8.7 Hz, 1H), 6.63 (d, $J$ = 8.3 Hz, 2H), 4.10 (dt, $J$ = 9.3, 9.4, 3.7 Hz, 1H), 3.77 (dq, $J$ = 17.4, 5.5 Hz, 2H), 2.88 (dd, $J$ = 14.1, 3.8 Hz, 1H), 2.58-2.64 (m, 1H), 1.30 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 172.2, 171.2, 155.6, 155.2, 130.0, 128.2, 114.7, 77.8, 55.8, 54.9, 40.6, 36.6, 28.1; IR (neat): ῡ = 3325, 2974, 2937, 1652, 1515, 1443, 1393, 1367, 1225, 1159, 1050, 1022, 828, 544 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{16}H_{21}N_2O_6$, 337.1; found, 337.1 $[M - H]^-$; $[\alpha]_D^{24}$ +0.29 (c 0.98, MeOH).

**Boc-D-Tyr-Gly-aminoethanol (15)**



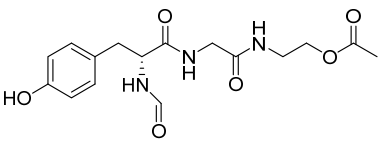

Boc-D-Tyr-Gly (**14**, 1.30 g, 3.85 mmol), DCC (0.95 g, 4.61 mmol) and HOBt (0.62 g, 4.61 mmol) were stirred in DMF (10.0 mL) at 25 °C for 30 minutes. Ethanolamine (0.23 mL, 3.85 mmol) was added and it was stirred for 12 hours. The reaction mixture was concentrated under reduced pressure and the residue dissolved in water. The dicyclohexyl urea was filtered off and the aqueous layer was extracted with DCM three times. The aqueous layer was then extracted with EtOAc (3x). Finally, the EtOAc phase was washed with 10 % $NaHCO_3$ solution and it was dried over $MgSO_4$. Evaporation of the solvent afforded the desired compound as a white solid in pure form (69 % yield). $^1$H NMR (500 MHz, $D_2O$-$d_6$): δ= 7.14 (d, $J$ = 8.0 Hz, 2H), 6.84 (d, $J$ = 8.1 Hz, 2H), 4.25 (t, $J$ = 7.9 Hz, 1H), 3.79-3.88 (m, 2H), 3.62 (t, $J$ = 5.4 Hz, 2H), 3.32 (t, J = 5.0 Hz, 2H), 2.04 (dd, $J$ = 14.0, 6.2 Hz, 1H), 2.85 (dd, $J$ = 16.0, 7.4 Hz, 1H), 1.34 (s, 9H); $^{13}$C NMR (125 MHz, $D_2O$-$d_6$): δ= 175.0, 171.3, 157.5, 154.3, 130.5, 128.3, 115.4, 81.6, 59.8, 56.5, 42.4, 41.4, 36.2, 27.4; IR (neat): ῡ = 3308, 2974, 2926, 1651, 1515, 1452, 1392, 1366, 1242, 1161, 1055, 828, 541, 424 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{18}H_{27}N_3NaO_6$, 404.1792; found, 404.1790 $[M + Na]^+$; $[\alpha]_D^{24}$ -16.3 (c 0.99, MeOH).

**D-Tyr-Gly-aminoethanol (16)**



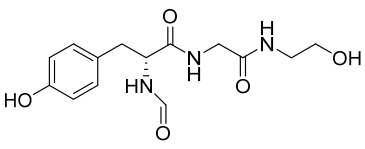
**15** (0.85 g, 2.22 mmol) was dissolved in EtOH/water 50:50 v/v (6.9 mL) and 20 % aqueous HCl solution (2.8 mL) was added dropwise. After stirring the reaction for 12 hours, it was neutralized with a 10 % NaOH solution. It was extracted with EtOAc and the organic phase was discarded. The aqueous phase was concentrated under reduced pressure, the residue was dissolved in MeOH, and the salts were filtered off to obtain the desired compound as a white solid (98 % yield). $^{1}$H NMR (400 MHz, $D_2O$-*d*$_6$): δ= 7.24 (d, *J* = 8.4 Hz, 2H), 6.96 (d, *J* = 8.5 Hz, 2H), 4.30 (t, *J* = 7.2 Hz, 1H), 3.95 (dd, J = 49.0, 16.7 Hz, 2H), 3.70 (t, *J* = 5.5 Hz, 2H), 3.39 (t, J = 5.5 Hz, 2H), 3.21 (d, *J* = 7.3 Hz, 2H); $^{13}$C NMR (100 MHz, $D_2O$-*d*$_6$): δ= 171.4, 170.4, 155.7, 131.4, 126.1, 116.4, 60.4, 55.1, 42.9, 42.0, 36.5. IR (neat): ῡ = 3263, 3069, 2931, 1652, 1539, 1514, 1444, 1227, 1051, 1017, 825, 549, 498 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{19}N_3NaO_4$, 304.1268; found, 304.1265 [M + Na]$^+$; $[\alpha]_D^{25}$ -42.6 (c 1.08, MeOH).

***N*-Formyl-D-Tyr-Gly-aminoethanyl acetate (17)**



To a solution of **16** (0.20 g, 0.712 mmol) in formic acid (4.0 mL), acetic anhydride was added dropwise at 0 °C and the solution was stirred at 25 °C for 12 hours. The reaction mixture was concentrated, dissolved in DCM and filtrated. The product was purified by flash column chromatography (DCM/10% MeOH) to obtain it in 41% yield as a white solid. $^{1}$H NMR (400 MHz, $CDCl_3$): δ= 8.11 (s, 1H), 8.04 (s, 0.6H), 7.20 (d, *J* = 8.4 Hz, 2H), 7.01 (d, *J* = 8.4 Hz, 2H), 6.95 (dd, *J* = 10.9, 5.8 Hz, 1H), 6.71 (t, *J* = 5.5 Hz, 0.6H), 6.61 (t, *J* = 8.1Hz, 1.4H), 4.70 (quint, *J* = 7.2, 7.0 Hz, 1H), 4.22-4.25 (m, 1H), 4.12-4.16 (m, 1H), 3.93 (ddd, *J* = 16.7, 5.9, 2.2 Hz, 1H), 3.76 (dd, *J* = 16.6, 5.5 Hz, 1H), 3.50 (dt, *J* = 11.0, 5.6 Hz, 2H), 3.09 (ddd, *J* = 29.3, 13.9, 7.1 Hz, 2H), 2.29 (s, 3H), 2.06 (s, 1.6H), presence of rotamers; $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 171.3, 170.9, 169.7, 168.8, 168.7, 161.6, 161.5, 161.1, 149.8, 133.6, 133.6, 130.3, 121.9, 62.8, 62.4, 53.5, 53.4, 43.1, 43.1, 38.8, 38.5, 37.1, 37.1, 21.1, 20.9; IR (neat): ῡ = 3272, 3083, 2932, 1760, 1720, 1637, 1551, 1508, 1438, 1390, 1371, 1220, 1201, 1182, 1165, 1054, 1018, 916, 719, 554 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{16}H_{21}N_3NaO_6$, 374.1323; found, 374.1321 [M + Na]$^+$; $[\alpha]_D^{26}$ -26.5 (c 0.36, MeOH).

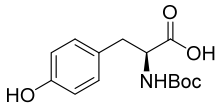
***N*-formyl-D-Tyr-Gly-aminoethanol (18)**



Ester **17** (125 mg, 0.405 mmol) was dissolved in MeOH (1.50 mL) and $K_2CO_3$ was added (0.1 g, 0.724mmol). The reaction mixture was stirred at 25 °C for 1 hour, it was filtrated and concentrated obtaining alcohol **18** in

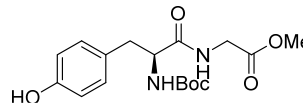
quantitative yield. $^{1}$H NMR (500 MHz, MeOD): δ= 8.04 (s, 1H), 6.91 (d, $J$ = 8.4 Hz, 2H), 6.58 (d, $J$ = 8.4 Hz, 2H), 4.44 (t, $J$ = 7.4 Hz, 1H), 3.76 (dd, $J$ = 46.6, 17.0 Hz, 2H), 3.56 (t, $J$ = 6.0 Hz, 2H), 3.35 (s, 2H), 2.87 (ddd, $J$ = 39.3, 13.7, 7.5 Hz, 2H); $^{13}$C NMR (500 MHz, MeOD): δ= 174.4, 171.7, 170.4, 170.4, 166.0, 164.0, 130.9, 122.8, 119.7, 61.4, 56.4, 43.7, 43.0, 38.0, 24.3, presence of rotamers; IR (neat): ῦ = 3272, 3078, 2932, 2807, 1760, 1719, 1638, 1596, 1557, 1499, 1439, 1389, 1372, 1221, 1166, 1020, 916, 844, 759, 718, 668, 634, 554, 531 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_{19}N_3NaO_5$, 332.1217; found, 332.1215 $[M + Na]^+$; $[\alpha]_D^{26}$ -16.6 (c 0.54, MeOH).

**Boc-L-Tyrosine (24)**



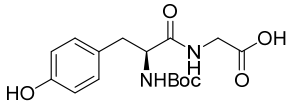

Following an identical procedure to the one described for compound **12**, starting from D-glutamic acid (3.00 g, 16.56 mmol), protected compound **24** was obtained as a white solid (4.04 g, 87 % yield). $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ= 12.53 (br, 1H), 9.20 (s, 1H), 7.02 (d, $J$ = 8.4 Hz, 2H), 6.65 (d, $J$ = 8.4 Hz, 2H), 3.98 (ddd, J = 10.0, 8.5, 4.6 Hz, 1H), 2.87 (dd, $J$ = 13.9, 4.6 Hz, 1H), 2.69 (dd, $J$ = 13.8, 10.2 Hz, 1H), 1.32 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 173.7, 155.8, 155.4, 129.9, 128.0, 114.8, 77.9, 55.5, 35.6, 28.1; IR (neat): ῦ = 3381, 3147, 1725, 1668, 1515, 1402, 1370, 1242, 1201, 1163, 1136, 1055, 822, 783, 654, 595, 533, 491, 456 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{19}NNaO_5$, 304.1; found, 304.1 $[M + Na]^+$; $[\alpha]_D^{23}$ +17.3 (c 1.06, MeOH).

**Boc-L-Tyr-D-Gly-OMe (25)**




Following an identical procedure to the one described for compound **13**, starting from Boc-L-tyrosine (2.00 g, 7.11 mmol) , ester **25** was obtained as a white solid (2.02 g, 80 % yield). $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ= 9.12 (s, 1H), 8.29 (t, $J$ = 5.5 Hz, 1H), 7.04 (d, $J$ = 8.1 Hz, 2H), 6.80 (d, $J$ = 8.7 Hz, 1H), 6.64 (d, $J$ = 8.1 Hz, 2H), 4.11 (dt, $J$ = 9.4, 9.1, 4.2 Hz, 1H), 3.79-3.92 (m, 2H), 3.63 (s, 3H), 2.87 (dd, $J$ = 13.9, 3.4 Hz, 1H), 2.58-2.67 (m, 1H), 1.30 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 172.4, 170.2, 155.6, 155.2, 130.0, 128.2, 114.7, 77.9, 55.8, 54.9, 51.6, 40.5, 36.6, 28.1; IR (neat): ῦ = 3366, 3332, 2979, 2952, 1688, 1636, 1515, 1439, 1358, 1223, 1164, 1048, 985, 812, 670, 564 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{17}H_{24}N_2NaO_6$, 375.2; found, 375.2 $[M + Na]^+$; $[\alpha]_D^{24}$ -0.29 (c 1.02, MeOH).

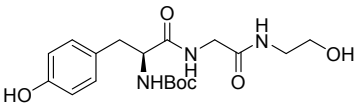
**Boc-L-Tyr-Gly (26)**




Following an identical procedure to the one described for compound **14**, starting from ester **25** (1.839 g, 5.221 mmol), acid **26** was obtained as a white solid (1.730 g, 98 % yield). $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ= 12.54 (br, 1H), 9.11 (s, 1H), 8.15 (t, $J$ = 5.7 Hz, 1H), 7.04 (d, $J$

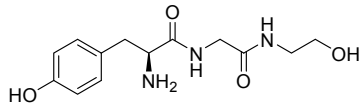
= 8.3 Hz, 2H), 6.78 (d, *J* = 8.8 Hz, 1H), 6.63 (d, *J* = 8.3 Hz, 2H), 4.10 (dt, *J* = 10.0, 3.9 Hz, 1H), 3.77 (dq, *J* = 17.6, 17.5, 5.9 Hz, 2H), 2.88 (dd, *J* = 13.9, 3.9 Hz, 1H), 2.58 – 2.64 (m, 1H), 1.30 (s, 9H); $^{13}$C NMR (125 MHz, DMSO-*d*$_6$): δ= 172.2, 171.2, 155.6, 155.2, 130.0, 128.2, 114.7, 77.9, 55.8, 40.6, 36.6, 28.1; IR (neat): ῦ = 3344, 2979, 2932, 1722, 1679, 1646, 1517, 1429, 1366, 1247, 1227, 1163, 1050, 822, 577, 547 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{16}H_{22}N_2O_6$, 337.1; found, 337.1 [M - H]$^-$.

**Boc-L-Tyr-Gly-aminoethanol (27)**



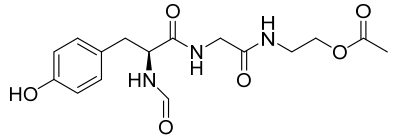
Following an identical procedure to the one described for compound **15**, starting from **Boc-L-Tyr-Gly** (0.200 g, 0.595 mmol), alcohol **27** was obtained as a white solid (0.186 g, 82% yield). $^1$H NMR (500 MHz, D$_2$O-*d*$_6$): δ= 7.21 (d, *J* = 7.9 Hz, 2H), 6.91 (d, *J* = 8.2 Hz, 2H), 4.33 (t, *J* = 7.5 Hz, 1H), 3.87-3.95 (m, 2H), 3.69 (t, *J* = 5.5 Hz, 2H), 3.39 (t, *J* = 5.2 Hz, 2H), 3.11 (dd, *J* = 13.7, 6.0 Hz, 1H), 2.92 (dd, *J* = 15.6, 7.1 Hz, 1H), 1.41 (s, 9H); $^{13}$C NMR (125 MHz, D$_2$O-*d*$_6$): δ= 175.0, 171.3, 157.5, 154.3, 130.5, 128.3, 115.4, 81.6, 59.8, 56.5, 42.4, 41.4, 36.2, 27.4; IR (neat): ῦ = 3296, 2974, 1651, 1515, 1451, 1366, 1244, 1161, 1055, 1019, 827, 543, 424, 416 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{18}H_{27}N_3NaO_6$, 404.2; found, 404.2 [M + Na]$^+$; $[\alpha]_D^{25}$ +17.8 (c 1.07, MeOH).

**L-Tyr-Gly-aminoethanol (28)**



Following an identical procedure to the one described for compound **16**, starting from **27** (0.186 g, 0.487 mmol), amine **28** was obtained in quantitative yield (0.185 g). $^1$H NMR (400 MHz, D$_2$O-*d*$_6$): δ= 7.28 (d, *J* = 8.4 Hz, 2H), 6.99 (d, *J* = 8.2 Hz, 2H), 4.36 (t, *J* = 7.2 Hz, 1H), 3.98 (dd, *J* = 49.0, 16.7 Hz, 2H), 3.73 (t, *J* = 5.5 Hz, 2H), 3.43 (t, *J* = 5.5 Hz, 2H), 3.24 (d, *J* = 7.2 Hz, 2H); $^{13}$C NMR (125 MHz, D$_2$O-*d*$_6$): δ= 170.9, 169.9, 155.1, 130.9, 125.5, 115.9, 59.9, 54.6, 42.5, 41.5, 35.9; IR (neat): ῦ = 3263, 3082, 2932, 2811, 2775, 2682, 1653, 1548, 1515, 1367, 1351, 1246, 1020, 826, 764, 553 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{13}H_{19}N_3NaO_4$, 304.1; found, 304.1 [M + Na]$^+$; $[\alpha]_D^{25}$ +20.3 (c 0.96, MeOH).

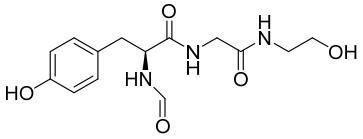
***N*-Formyl-L-Tyr-Gly-aminoethanyl acetate (29)**



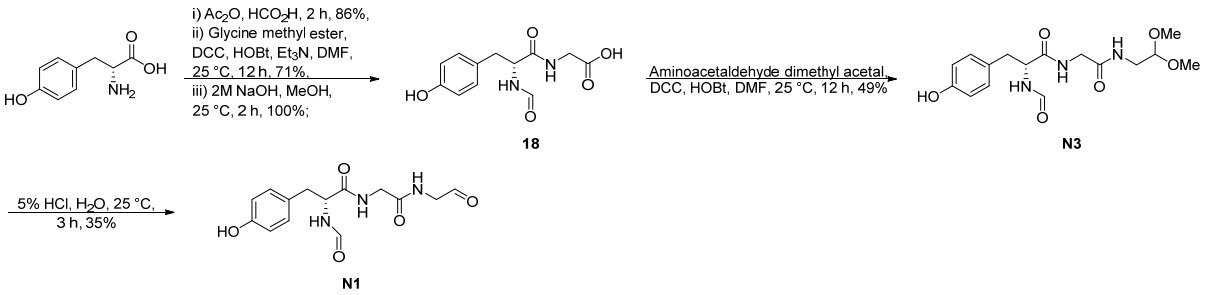
Following an identical procedure to the one described for compound **17**, starting from L-Tyr-Gly-aminoethanol (200 mg, 0.712 mmol), formamide **29** was obtained (108 mg, 43 % yield). $^1$H NMR (500 MHz, CDCl$_3$): δ= 8.00 (d, *J* = 4.1 Hz, 1.5H), 7.64 (t, *J* = 5.5 Hz, 0.53H), 7.60 (t, J = 5.4Hz, 0.33H), 7.29 (d, *J* = 7.3 Hz, 0.56H), 7.24 (d, *J* = 7.4 Hz, 0.35H), 7.17 (d, *J* = 8.5 Hz, 2H), 7.14 (t, *J* = 5.6Hz, 0.53H), 7.03 (t, *J* = 5.6Hz, 0.36H), 6.96 (d, *J* = 8.5 Hz, 2H), 4.73-4.79 (m, 1H), 4.19 (t, *J* = 5.5 Hz, 1.20H), 4.10 (t, *J* = 5.7

Hz, 0.71H), 3.84-3.89 (m, 1H), 3.76 (dd, $J$ = 16.7, 5.4 Hz, 1H), 3.45 (td, $J$ = 19.3, 5.6 Hz, 2H), 3.08-3.12 (m, 1H), 2.99 (dd, $J$ = 14.0, 7.6 Hz, 2H), 2.26 (s, 3H). presence of rotamers; $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 171.3/171.3/171.2, 169.9/169.9, 169.2, 169.1, 162.0, 161.2, 149.5, 133.9, 130.3, 121.7, 62.7, 62.2, 53.3, 53.2, 42.9, 38.5, 38.3, 37.2, 37.1, 21.1, 20.9, presence of rotamers; IR (neat): ῡ = 3308, 2937, 1764, 1723, 1650, 1507, 1370, 1220, 1193, 1167, 1018, 913, 545 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{16}H_{21}N_3NaO_6$, 374.1; found, 374.3 [M + Na]$^+$; $[\alpha]_D^{25}$ -4.44 (c 1.02, MeOH).

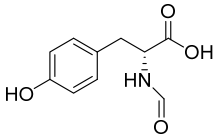
### *N*-Formyl-L-Tyr-Gly-aminoethanol (N5)



**29** (50 mg, 0.143 mmol) was dissolved in MeOH (0.6 mL) and $K_2CO_3$ (40 mg, 0.289 mmol) was added. It was stirred at 25 °C until dissapearance of the starting material. The reaction mixture was filtrated and concentrated. The residue was dissolved in MeOH and $K_2CO_3$ was filtered off to obtain **N5** in quantitative yield. $^1$H NMR (500 MHz, $D_2O$): δ= 7.96 (s, 1H), 6.94 (d, $J$ = 8.5 Hz, 2H), 6.55 (d, J = 8.5 Hz, 2H), 4.46 (t, $J$ = 7.2 Hz, 1H), 3.83 (dd, $J$ = 32.4, 17.0 Hz, 2H), 3.62 (t, $J$ = 5.5Hz, 2H), 3.52 (t, $J$ = 5.5 Hz, 2H), 3.22 (t, $J$ =5.5 Hz, 2H), 2.87 (dd, $J$ = 7.4, 2.1 Hz, 2H); $^{13}$C NMR (125 MHz, $D_2O$): δ= 173.9, 171.2, 170.9, 165.6, 164.1, 162.6, 130.3, 122.5, 117.9, 59.7, 54.4, 42.4, 41.3, 36.0, presence of rotamers; IR (neat): ῡ = 3272, 2941, 2822, 1651, 1597, 1496, 1441, 1380, 1269, 1169, 1076, 1026, 829, 554 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{19}N_3NaO_5$, 332.1; found, 332.1 [M + Na]$^+$; $[\alpha]_D^{25}$ +0.17 (c 1.40, MeOH).



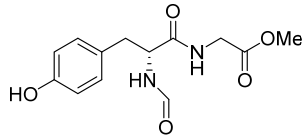

### *N*-Formyl-D-Tyr (30)



D-Tyrosine (1.50 g, 0.83 mmol) was dissolved in formic acid (9.8 mL) and acetic anhydride (5.0 mL, 52.9 mmol) was added in excess. The reaction mixture was stirred at 10 °C for 1 hour and then at 25 °C for 30 minutes. Water (6.50 mL) was added at 10 °C, it was stirred for 30 minutes and the reaction mixture was then concentrated under reduced pressure. The residue was dissolved in 1M HCl solution and extracted with EtOAc. The combined organic phases were dried over $MgSO_4$ to afford acid **18** as a white solid in 85 % yield. $^1$H NMR (500 MHz, DMSO-$d_6$): δ= 12.77 (br, 1H), 9.23 (s, 1H), 8.31 (d, $J$ = 8.0 Hz, 1H), 7.96 (s, 1H), 6.99 (d, $J$ = 8.4 Hz, 2H), 6.65 (d, $J$ = 8.5 Hz, 2H), 4.42 (dt, $J$ = 8.5, 5.1 Hz, 1H), 2.94 (dd, $J$ = 13.9, 4.9

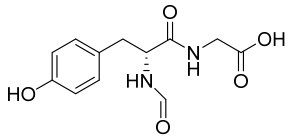
Hz, 1H), 2.74 (dd, $J$ = 13.9, 9.0 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$):δ= 172.6, 160.9, 155.9, 129.0, 127.1, 114.9, 52.1, 35.6; IR (neat): ῡ = 3336, 3058, 3025, 2919, 2815, 1713, 1625, 1611, 1601, 1536, 1513, 1445, 1427, 1376, 1344, 1272, 1221, 1173, 1109, 1071, 950, 828, 790, 661, 546, 471 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{10}H_{10}NO_4$, 208.0; found, 207.9 [M - H]$^-$; $[\alpha]_D^{24}$ -49.6 (c 1.22, MeOH).

### *N*-Formyl-D-Tyr-Gly-OMe (31)



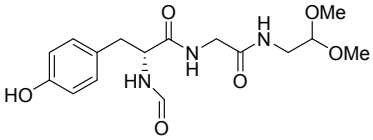

N-formyl-D-tyrosine (**18**, 1.30 g, 6.24 mmol) was dissolved in DMF (25.0 mL). HOBt (1.01 g, 7.49 mmol) and DCC (1.55 g, 7.49 mmol) were added and the reaction mixture was stirred at 25 °C for 30 minutes. Glycine methyl ester (1.57 g, 12.50 mmol) and triethyl amine (1.0 mL, 7.49 mmol) were added and it was stirred for 12 hours at 25 °C. The reaction mixture was then filtered and concentrated. The desired coupled peptide was purified by flash column chromatography (acetonitrile/water 1:0.05) to obtain it as a white solid in 71 % yield. $^1$H NMR (500 MHz, MeOH): δ= 8.01 (s, 1H), 7.07 (d, $J$ = 8.4 Hz, 2H), 6.70 (d, $J$ = 8.4 Hz, 2H), 4.68 (dd, $J$ = 8.3, 5.7 Hz, 1H), 3.89-3.96 (m, 2H), 3.72 (s, 3H), 3.08 (dd, $J$ = 14.0, 5.8 Hz, 1H), 2.82 (dd, $J$ = 13.9, 8.6 Hz, 1H); $^{13}$C NMR (125 MHz, MeOH): δ= 174.3, 173.8, 171.5, 163.5, 157.5, 157.4, 131.4, 128.7, 128.6, 116.2, 54.7, 53.9, 52.7, 41.9, 38.4, 37.7, presence of rotamers; IR (neat): ῡ = 3293, 3161, 2954, 2916, 1742, 1633, 1591, 1564, 1539, 1516, 1435, 1389, 1354, 1219, 1172, 1113, 1053, 1026, 984, 909, 830, 752, 723, 680, 560, 472 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{13}H_{16}N_2NaO_5$, 303.0951; found, 303.0948 [M + Na]$^+$;$[\alpha]_D^{25}$ -6.1 (c 1.09, MeOH).

### *N*-Formyl-D-Tyr-Gly (18)



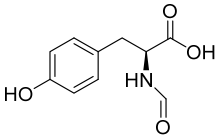

N-formyl-D-Tyr-Gly-OMe (**31**, 1.00 g, 3.58 mmol) was dissolved in MeOH (16.0 mL) and a 2M NaOH solution (3.7 mL) was added. The reaction mixture was stirred at 25 °C for 2.5 hours. 1M HCl was added to neutralize the reaction mixture and it was extracted with DCM. The organic phases were discarded and the aqueous phase was concentrated. The residue was dissolved in MeOH and the salts were filtered out. The desired acid was obtain in quantitative yield as a white solid after evaporating the MeOH. $^1$H NMR (400 MHz, MeOH): δ= 8.29 (t, $J$ = 5.9 Hz, 1H), 8.03 (s, 1H), 7.07 (d, $J$ = 8.4 Hz, 2H), 6.71 (d, $J$ = 8.5 Hz, 2H), 4.67 (dd, $J$ = 8.7, 5.5 Hz, 1H), 3.87 (d, $J$ = 5.5 Hz, 2H), 3.10 (dd, J = 14.1, 5.3 Hz, 1H), 2.83 (dd, $J$ = 14.0, 8.9 Hz, 1H); $^{13}$C NMR (100 MHz, MeOH): δ= 173.7, 172.7, 163.5, 157.4, 131.4, 128.8, 116.3, 54.7, 41.8, 38.4; IR (neat): ῡ = 3326, 3031, 2926, 2822, 1658, 1613, 1514, 1447, 1384, 1235, 1174, 1101, 1021, 830, 782, 741, 551, 516 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{12}H_{14}N_2NaO_5$, 289.0795; found, 289.0793 [M + Na]$^+$; $[\alpha]_D^{25}$ -6.8 (c 1.05, MeOH).

**_N_-Formyl-D-Tyr-Gly-aminoacetaldehyde (N3)**



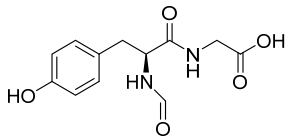
N-formyl-D-Tyr-Gly (**18**, 0.50 g, 1.89 mmol) was dissolved in DMF (5.0 mL). HOBt (0.51 g, 3.78 mmol) and DCC (0.78 g, 3.78 mmol) were added and stirred at 25 °C for 30 minutes. Acetaldehyde (0.41 mL, 3.78 mmol) was then added and the reaction mixture was stirred for 12 hours. The reaction mixture was concentrated and the desired acetaldehyde was purified by flash column chromatography (acetonitrile/water 1:0.05). The product was obtained as a white solid in 49 % yield (0.29 g, 0.82 mmol). $^{1}$H NMR (500 MHz, MeOH): δ= 8.03 (s, 1H), 7.06 (d, $J$ = 8.6 Hz, 2H), 6.71 (d, $J$ = 8.6 Hz, 2H), 4.53 (t, $J$ = 7.4 Hz, 1H), 4.42 (t, $J$ = 5.4 Hz, 1H), 3.79 (dd, $J$ = 74.8, 16.8 Hz, 2H), 3.36 (d, $J$ = 0.8 Hz, 6H), 3.05 (dd, $J$ = 13.9, 6.5 Hz, 2H), 2.85 (dd, $J$ = 13.9, 8.3 Hz, 2H); $^{13}$C NMR (100 MHz, MeOH): δ= 173.9, 171.6, 163.9, 157.5, 131.3, 128.6, 116.3, 97.1, 55.6, 45.5, 43.5, 37.9, 1 carbon is missing due to overlapping; IR (neat): $\tilde{\upsilon}$ = 3285, 3088, 2943, 2838, 1633, 1557, 1541, 1515, 1392, 1234, 1219, 1136, 1106, 1069, 1025, 835, 722, 697, 683, 555, 413 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{16}H_{23}N_3NaO_6$, 376.1479; found, 376.1480 $[M + Na]^+$; $[\alpha]_D^{25}$ -26.8 (c 1.16, MeOH).

**_N_-Formyl-L-Tyr (32)**



Following an identical procedure to the one described for compound **30**, starting from L-tyrosine (2.00 g, 11.04 mmol), formamide **32** was obtained as a white solid (1.91 g, 83% yield). $^{1}$H NMR (500 MHz, DMSO-$d_6$): δ= 12.81 (br, 1H), 9.23 (s, 1H), 8.31 (d, $J$ = 8.4 Hz, 1H), 7.96 (s, 1H), 6.99 (d, $J$ = 8.3 Hz, 2H), 6.65 (d, $J$ = 8.4 Hz, 2H), 4.42 (dt, $J$ = 8.6, 5.4 Hz, 1H), 2.94 (dd, $J$ = 13.9, 4.8 Hz, 1H), 2.74 (dd, $J$ = 13.8, 8.9 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$):δ = 172.6, 160.9, 155.9, 130.0, 127.2, 114.9, 52.2, 35.9; IR (neat): $\tilde{\upsilon}$ = 3335, 3067, 3026, 2922, 2807, 1713, 1625, 1611, 1601, 1535, 1513, 1445, 1427, 1376, 1345, 1271, 1221, 1173, 1109, 1071, 950, 828, 790, 661, 546, 470 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{10}H_{10}NO_4$, 208.0; found, 207.9 $[M - H]^-$; $[\alpha]_D^{24}$ +50.3 (c 0.97, MeOH).

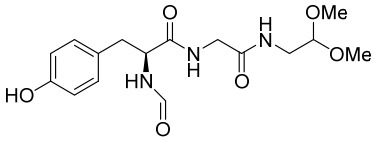
**_N_-Formyl-L-Tyr-Gly (33)**



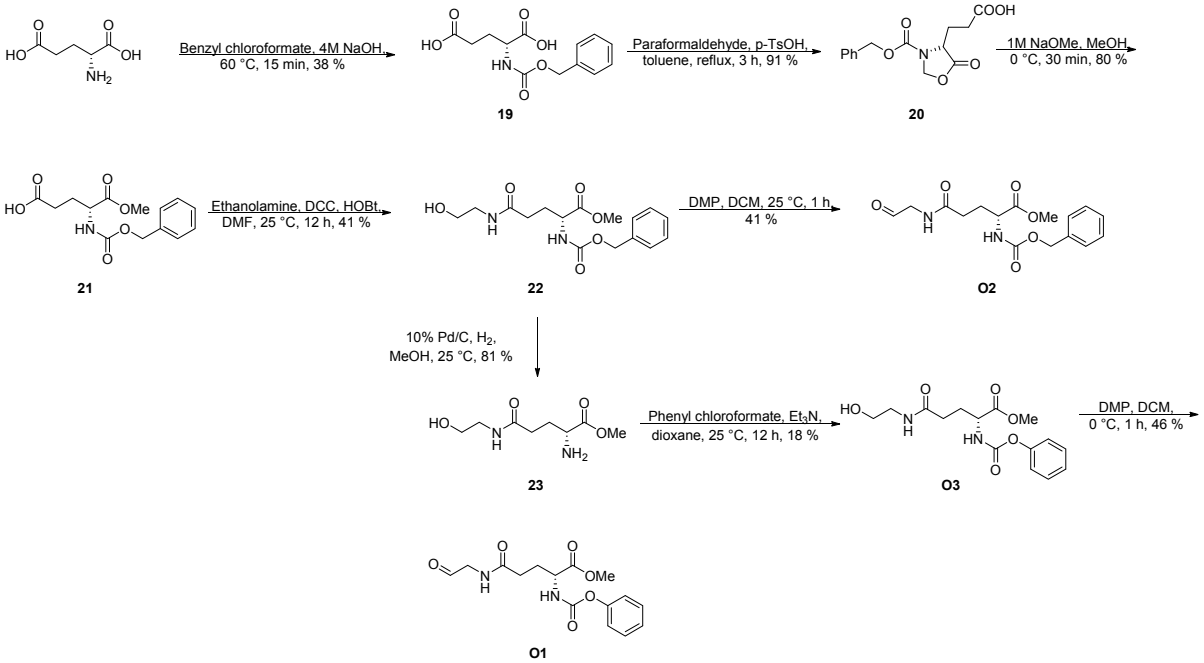
Following an identical procedure to the one described for compound **31**, starting from **32** (1.00 g, 3.58 mmol), acid **33** was obtained as a white solid (80 % yield over two steps). $^{1}$H NMR (500 MHz, MeOH): δ= 8.02 (s, 1H), 7.66-7.61 (m, 1H), 7.25-7.31 (m, 1H), 7.06 (d, $J$ = 8.3 Hz, 2H), 6.70 (d, $J$ = 8.3 Hz, 2H), 4.66 (dd, $J$ = 8.8, 5.3 Hz, 1H), 3.75 (dd, $J$ = 55.5, 17.3 Hz, 2H), 3.11 (dd, J = 14.1, 5.2 Hz, 1H), 2.82 (dd, $J$ = 14.0, 9.1 Hz, 1H); $^{13}$C NMR (125 MHz, MeOH): δ= 175.7, 172.9, 163.6, 157.3, 131.4, 129.0, 116.2, 55.0, 44.3, 38.2; IR (neat): $\tilde{\upsilon}$ = 3274, 2939, 2826, 1660, 1596, 1515, 1447, 1385, 1241, 1174, 1102, 1022, 832, 782, 743,

549, 516 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{12}H_{13}N_2O_5$, 265.1; found, 264.9 $[M - H]^-$; $[\alpha]_D^{25}$ +6.9 (c 1.07, MeOH).

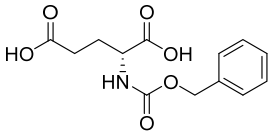
### *N*-Formyl-L-Tyr-Gly-aminoacetaldehyde (N6)



Following an identical procedure to the one described for compound **N3**, starting from acid **33** (0.50 g, 1.89 mmol), acetaldehyde **N6** was obtained as a white solid (0.11 g, 17 % yield). $^1$H NMR (500 MHz, MeOH): δ= 8.04 (s, 1H), 7.67 (t, *J* = 7.7 Hz, 1H), 7.07 (d, *J* = 8.5 Hz, 2H), 6.72 (d, *J* = 8.5 Hz, 2H), 4.57 (t, *J* = 7.4 Hz, 1H), 4.42 (t, *J* = 5.4 Hz, 1H), 3.81 (ddd, *J* = 67.5, 16.8 Hz, 4.5 Hz, 2H), 3.36 (s, 6H), 3.06 (dd, *J* = 13.9, 6.4 Hz, 2H), 2.87 (dd, *J* = 13.9, 8.3 Hz, 2H); $^{13}$C NMR (100 MHz, MeOH): δ= 173.9, 171.637, 163.9, 157.4, 131.3, 128.6, 116.3, 103.7, 55.5, 43.5, 42.2, 42.1, 37.9; IR (neat): ῡ = 3343, 2929, 2943, 1645, 1613, 1515, 1467, 1343, 1257, 1228, 1175, 1103, 1061, 980, 905, 811, 665, 487, 440, 421 $cm^{-1}$. MS (ESI), *m/z*: calcd for $C_{16}H_{23}N_3NaO_6$, 376.1; found, 376.1 $[M + Na]^+$; $[\alpha]_D^{25}$ +10.2 (c 1.09, MeOH).



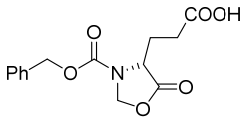

### *N*-Benzyloxycarbonyl-*D*-glutamic acid (19)[95]



To a solution of L-glutamic acid (5.00 g, 33.98 mmol) in 17.0 mL of 4M NaOH, benzyl chloroformate (7.76 mL, 54.37 mmol) was added dropwise at 0 °C. The reaction mixture was heated at 60 °C for 15 minutes and cooled to 25 °C. It was extracted with diethyl ether, the aqueous phase was acidified to pH 2 by 1M HCl addition and extracted with EtOAc (3x). The combined EtOAc phases were dried with $MgSO_4$ to obtained the protected glutamic acid in 38 % yield. $^1$H NMR (500 MHz, DMSO-*d*$_6$): δ= 12.25 (br, 2H), 7.47 (d, *J* = 7.4 Hz, 1H), 7.31-7.38 (m, 5H), 5.04 (s, 2H), 4.02 (dt, *J* = 8.9, 5.1 Hz, 1H), 2.30 (dd, *J* =14.1, 6.6 Hz, 2H),

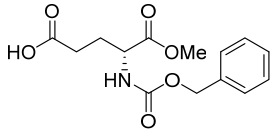
1.98 (dq, $J$ = 7.6, 5.4 Hz, 1H), 1.79 (dq, $J$ = 9.0, 6.4 Hz, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 173.3, 136.8, 128.0, 127.5, 127.4, 65.2, 52.9, 29.9, 26.0, two carbons are missing due to overlapping; IR (neat): $\tilde{\upsilon}$ = 3303, 3037, 2945, 1700, 1685, 1547, 1525, 1418, 1256, 1169, 1051, 1014, 951, 910, 869, 779, 734, 694, 649, 599, 573 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{13}H_{15}NNaO_6$, 304.1; found, 304.1 $[M + Na]^+$; $[\alpha]_D^{24}$ + 11.8 (c 0.65, MeOH).

**4-(R)-(2-Carboxy-ethyl)-5-oxo-oxazolidine-3-carboxylic acid benzyl ester (20)**[96]



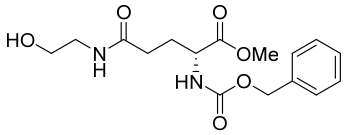

To **19** (3.63 g, 12.92 mmol) in toluene (85.0 mL), 1.5 equivalents of paraformaldehyde and 1% of toluensulfonic acid (wt/wt) were added. The reaction mixture was refluxed under Dean-Stark conditions for 3 hours. The mixture was diluted with EtOAc, washed with brine and dried over $MgSO_4$. The solvent was evaporated under reduced pressure to give the azalactone **20** in 91 % yield, which was used without further purification. $^1$H NMR (500 MHz, DMSO-$d_6$): δ= 12.24 (br, 1H), 7.32-7.41 (m, 5H), 5.44 (d, $J$ =2.8 Hz, 1H), 5.29 (s, 1H), 5.12-5.18 (m, 2H), 4.38-4.45 (m, 1H), 2.22-2.39 (m, 2H), 2.09-2.19 (m, 2H), 1.96-2.06 (m, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$): δ= 173.4, 172.3, 136.0, 128.4, 128.0/128.0, 127.7/127.6, 77.8/77.7, 66.7/66.7, 53.9, 51.3, 28.9, 28.7, presence of rotamers; IR (neat): $\tilde{\upsilon}$ = 3304, 3026, 2951, 1799, 1702, 1546, 1525, 1417, 1356, 1254, 1167, 1131, 1051, 1015, 951, 913, 752, 735, 696 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_{15}NNaO_6$, 316.0792; found, 316.0790 $[M + Na]^+$; $[\alpha]_D^{29}$ -38.3 (c 1.07, MeOH).

**1-Methyl-*N*-Cbz-D-glutamic ester (21)**[97]



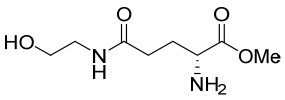

Azalactone **20** (1.21g, 4.11 mmol) was dissolved in dried MeOH (96.0 mL) and 5.8 mL of an *in situ* prepared NaOMe 1.75M solution (2eq.) was added at 0 °C. The reaction was run for 30 minutes at 0 °C. 1M HCl solution was added to the reaction mixture in order to acidify the reaction to pH 2 and it was extracted with EtOAc. The combined organic phases were dried over $MgSO_4$ and concentrated under reduced pressure to obtain the monomethyl ester **21** in 80 % yield. $^1$H NMR (500 MHz, DMSO-$d_6$): δ= 12.22 (br, 1H), 7.78 (t, $J$ = 6.3Hz, 1H), 7.26-7.39 (m, 5H), 5.03 (s, 2H), 4.09 (dd, $J$ = 13.3, 8.1 Hz, 1H), 2.37-2.43 (m, 1H), 2.26-2.34 (m, 1H), 1.92-2.03 (m, 1H), 1.73-1.85 (m, 1H); $^{13}$C NMR (125 MHz, DMSO-$d_6$) (2 rotamers): δ= 173.6, 172.4, 156.1, 136.8, 128.4, 128.3, 127.8, 127.7, 127.6, 65.5/65.5, 53.0/52.9, 51.9/51.9, 51.3, 29.8/29.6, 25.9/25.8; IR (neat): $\tilde{\upsilon}$ = 3303, 3032, 2956, 1701, 1524, 1436, 1418, 1345, 1256, 1212, 1178, 1050, 1015, 951, 912, 778, 735, 696, 599 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{14}H_{17}NNaO_6$, 318.0948; found, 318.0945 $[M + Na]^+$; $[\alpha]_D^{25}$ +22.7 (c 1.05, MeOH).

**1-Methyl-4-hydroxyethylcarbamoyl-*N*-Cbz-D-glutamic ester (22)**



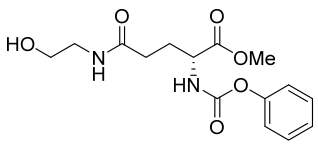
0.90 g (3.05 mmol) of methyl ester **21**, DCC (0.76 g, 3.66 mmol) and HOBt (0.49 g, 3.66 mmol) were dissolved in DMF (8.0 mL) at 25 °C. Ethanolamine (0.19 g, 3.05 mmol) was added after 30 minutes and the reaction was stirred for 12 hours. The reaction mixture was filtered, concentrated and the desired product was purified by flash column chromatography (DCM/5 % MeOH) to get it as a white solid in 41% yield. $^{1}$H NMR (400 MHz, $CDCl_3$): δ= 7.29-7.37 (m, 5H), 6.42 (s, 1H), 5.84 (d, $J$ = 7.4 Hz, 1H), 5.09 (s, 2H), 4.34 (t, $J$ = 6.9 Hz, 1H), 3.73 (s, 3H), 3.61-3.69 (m, 2H), 3.41-3.48 (m, 1H), 3.23-3.33 (m, 1H), 2.89 (br, 1H), 2.19-2.30 (m, 3H), 1.88-1.97 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 172.8, 172.5, 156.4, 136.0, 128.5, 128.2, 128.1, 67.1, 61.7, 53.3, 52.6, 42.4, 32.2, 28.5; IR (film): $\tilde{\upsilon}$ = 3308, 3067, 3031, 2943, 1707, 1650, 1536, 1454, 1438, 1339, 1262, 1216, 1177, 1057, 746, 698 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{16}H_{22}N_2NaO_6$, 361.1370; found, 361.1369 $[M + Na]^+$; $[\alpha]_D^{26}$ +13.3 (c 1.00, MeOH).

**1-Methyl-4-hydroxyethylcarbamoyl -D-glutamic ester (23)**



Alcohol **22** (0.199 g, 0.561 mmol) was dissolved in dried MeOH (2.0 mL). 10% Pd on activated charcoal (19.9 mg) were added and the reaction was carried out under a $H_2$ atmosphere at 25 °C until disappearance of the starting material. The reaction mixture was filtered through celite and washed with MeOH to obtain **30** (0.139 g, 81% yield), which was used without further purification. $^{1}$H NMR (500 MHz, MeOD): δ= 3.73 (s, 3H), 3.59 (t, $J$ = 5.8 Hz, 2H), 3.52 (dd, $J$ = 7.2, 5.9 Hz, 1H), 3.29 (t, $J$ =5.8 Hz, 2H), 2.32 (ddd, $J$ = 8.1, 4.8, 1.8 Hz, 2H), 2.00-2.07 (m, 1H), 1.85-1.92 (m, 1H); $^{13}$C NMR (125 MHz, MeOD): δ= 176.3, 175.3, 61.6, 54.6, 52.7, 43.0, 33.0, 31.1; IR (film): $\tilde{\upsilon}$ = 3288, 3078, 2936, 2865, 1732, 1644, 1551, 1438, 1380, 1210, 1176, 1065 cm$^{-1}$; HRMS (ESI), *m/z*: calcd for $C_8H_{17}N_2O_4$, 205.1183; found, 205.1184 $[M + H]^+$; $[\alpha]_D^{26}$ -10.5. (c 0.76, MeOH).

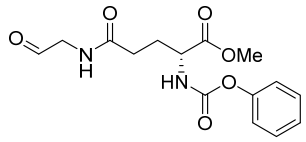
**Phenyl (R)-3-(2-hydroxyethylcarbamoyl)-1(methoxycarbonyl)propylcarbamate (O3)**



Amine **30** (105 mg, 0.51 mmol) was dissolved in dioxane (2.0 mL) and $Et_3N$ was added (143 μL). The solution was cooled to 0 °C and phenyl chloroformate (71 μL, 0.51 mmol) was added dropwise. The reaction was run at 25 °C for 12 hours. It was diluted with water and extracted with EtOAc. The residue was purified by flash column chromatography (DCM/5 % MeOH) obtaining the desired product in pure form (32.1mg, 18% yield). $^{1}$H NMR (500 MHz, $CDCl_3$): δ= 7.34 (t, $J$ = 7.6 Hz, 2H), 7.19 (t, $J$ = 7.4 Hz, 1H), 7.10 (d, $J$ = 7.8 Hz, 2H), 6.60 (t, $J$ = 5.5 Hz, 1H), 6.43 (d, $J$ = 8.0 Hz, 1H), 4.38 (dt,

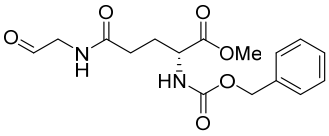
$J$ = 8.5, 4.1 Hz, 1H), 3.75 (s, 3H), 3.59-3.67 (m, 2H), 3.36-3.42 (m, 1H), 3.25-3.31 (m, 1H), 2.22-2.33 (m, 3H), 1.97-2.05 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 172.9, 172.3, 154.8, 150.7, 129.3, 125.5, 121.5, 61.6, 53.6, 52.7, 42.3, 32.1, 28.1; IR (film): $\tilde{\upsilon}$ = 3323, 3052, 2948, 1726, 1650, 1538, 1490, 1442, 1354, 1207, 1066, 1025, 688, 482, 418 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{15}H_{20}N_2NaO_6$, 347.1214; found, 347.1216 $[M + Na]^+$; $[\alpha]_D^{26}$ + 0.36 (c 0.65, MeOH).

**Phenyl-(R)-3-(formylmethylcarbamoyl)-1-(methoxycarbonyl)propylcarbamate (O1)**



To a solution of alcohol **O3** (21.3 mg, 0.060 mmol) in DCM (0.6 mL), DMP (30.6 mg, 0.072 mmol) was added at 0 °C. The reaction mixture was stirred for 15 minutes at 0 °C and at 25 °C for 2 hours. The reaction was quenched with an aqueous basic $Na_2S_2O_3$ solution in saturated sodium bicarbonate, it was stirred for 15 minutes and extracted with EtOAc. The combined organic phases were dried over $MgSO_4$, concentrated, and the desired product **O1** was purified by flash column chromatography (EtOAc). (10.0 mg, 45% yield). $^1$H NMR (500 MHz, $CDCl_3$): δ= 9.64 (s, 1H), 7.36 (t, $J$ = 7.9 Hz, 2H), 7.20 (t, $J$ = 7.4 Hz, 1H), 7.12 (t, $J$ = 7.9 Hz, 2H), 6.62 (br, 1H), 6.07 (d, $J$ = 7.7 Hz, 1H), 4.49 (dt, $J$ = 8.4, 4.2 Hz, 1H), 4.19 (ddd, $J$ = 39.7, 20.4, 5.1 Hz, 2H), 3.79 (s, 3H), 2.44 (dd, $J$ = 13.3, 6.6 Hz, 1H), 2.28-2.34 (m, 1H), 2.03-2.10 (m, 1H), 1.23-1.27 (m, 1H), 0.83-0.89 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 196.4, 172.3, 172.2, 154.7, 150.7, 129.3, 125.6 121.5 53.5 52.7, 50.3, 32.0, 28.6; IR (film): $\tilde{\upsilon}$ = 3327, 3062, 2958, 2922, 2859, 1732, 1654, 1530, 1489, 1438, 1262, 1205, 1025, 734, 690, 476, 438, 406 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{15}H_{18}N_2NaO_6$, 345.1057; found, 345.1052 $[M + Na]^+$;$[\alpha]_D^{25}$ + 2.90 (c 0.55, DCM).

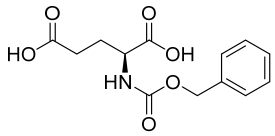
**1-Methyl-4-formylmethylcarbamoyl-N-Cbz-D-glutamic ester (O2)**



Methyl ester **22** (48.5 mg, 0.14 mmol) was dissolved in 1.0 mL of non-dried DCM. It was cooled to 0 °C and DMP (0.072 g) was added. The reaction mixture was stirred for 15 minutes at 0 °C and at 25 °C for 1 hour. The reaction was quenched with an aqueous basic $Na_2S_2O_3$ solution in saturated sodium bicarbonate, it was stirred for 15 minutes and extracted with EtOAc. The combined organic phases were dried over $MgSO_4$, concentrated, and the desired product **O2** was purified by flash column chromatography (EtOAc). (19.8 mg, 41% yield). $^1$H NMR (500 MHz, $CDCl_3$): δ= 9.63 (s, 1H), 7.30-7.36 (m, 5H), 6.54 (br, 1H), 5.60 (d, $J$ = 7.0 Hz, 1H), 5.07-5.14 (m, 2H), 4.44 (dt, $J$ = 8.6, 8.2, 4.0 Hz, 1H), 4.17 (dq, $J$ = 20.4, 20.3, 4.9 Hz, 2H), 3.75 (s, 3H), 2.32-2.41 (m, 2H), 2.21-2.30 (m, 1H), 1.93-2.02 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 196.3, 172.3, 172.2, 156.3, 136.1, 128.6, 128.3, 128.1, 67.2, 53.3, 52.6, 50.9, 50.3, 32.0, 28.7; IR (film): $\tilde{\upsilon}$ = 3307, 3062,

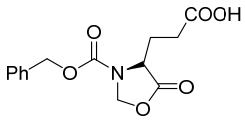
3031, 2952, 1716, 1655, 1530, 1454, 1438, 1260, 1214, 1053, 739, 699, 411 $cm^{-1}$; HRMS (ESI), *m/z*: calcd for $C_{16}H_{20}N_2NaO_6$, 359.1214; found, 359.1210 $[M + Na]^+$; $[\alpha]_D^{25}$ -4.38 (c 0.34, MeOH).

**_N_-Benzyloxycarbonyl-_L_-glutamic acid (34)**



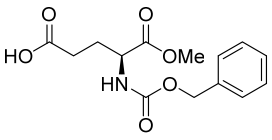
Following an identical procedure to the one described for compound **19**, starting from L-glutamic acid (5.00 g, 33.98 mmol), diacid **34** was obtained as a white solid (3.69 g, 38 % yield). $^1H$ NMR (400 MHz, DMSO-$d_6$): δ= 12.37 (br, 2H), 7.58 (d, $J$ = 8.1 Hz, 1H), 7.29-7.36 (m, 5H), 5.03 (s, 2H), 4.00 (dt, $J$ = 9.4, 5.0 Hz, 1H), 2.28-2.32 (m, 2H), 1.97 (dq, $J$ = 7.6, 7.7, 5.0 Hz, 1H), 1.71-1.81 (m, 1H); $^{13}C$ NMR (100 MHz, DMSO-$d_6$): δ= 173.6, 173.5, 156.1, 136.9, 128.3, 127.7, 127.6, 65.3, 53.0, 30.0, 26.0; IR (neat): $\tilde{\upsilon}$ = 3302, 3037, 2945, 1701, 1685, 1547, 1525, 1418, 1257, 1169, 1051, 1015, 951, 909, 779, 734, 694, 651, 599 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{13}H_{14}NO_6$, 280.1; found, 279.9 $[M - H]^-$; $[\alpha]_D^{25}$ = - 16.3 (c = 1.01 in MeOH).

**4-(S)-(2-Carboxy-ethyl)-5-oxo-oxazolidine-3-carboxylic acid benzyl ester (35)**



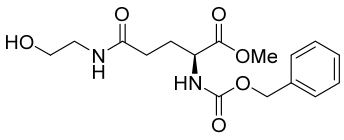
Following an identical procedure as the one described for compound **20**, starting from diacid **34** (3.40 g, 12.09 mmol), azalactone **35** was obtained (2.86 g, 91% yield). $^1H$ NMR (400 MHz, DMSO-$d_6$): δ= 12.21 (br, 1H), 7.31-7.41 (m, 5H), 5.45 (d, $J$ = 4.0 Hz, 1H), 5.29 (s, 1H), 5.12-5.18 (m, 2H), 4.41 (t, $J$ = 5.6 Hz, 1H), 2.21-2.38 (m, 2H), 2.09-2.17 (m, 1H), 1.96-2.05 (m, 1H); $^{13}C$ NMR (125 MHz, DMSO-$d_6$) : δ= 173.4, 172.4, 136.1, 128.4, 128.0, 127.6, 77.8, 66.7, 53.9, 48.6, 28.9, 25.6; IR (neat): $\tilde{\upsilon}$ = 3299, 3032, 2951, 1798, 1702, 1524, 1416, 1358, 1252, 1212, 1166, 1130, 1051, 1015, 916, 750, 735, 696, 614, 568, 493 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{14}NO_6$, 292.1; found, 292.0 $[M - H]^-$; $[\alpha]_D^{25}$ = + 41.6 (c = 1.04 in MeOH).

**1-Methyl-_N_-Cbz-L-glutamic ester (36)**



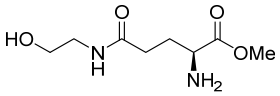
Following an identical procedure as the one described for compound **21**, starting from acid **35** (0.75 g, 2.55 mmol), mono-ester **36** was obtained (0.64 g, 85% yield). $^1H$ NMR (300 MHz, DMSO-$d_6$): δ= 11.99 (br, 1H), 7.77 (d, $J$ = 7.7 Hz, 1H), 7.31-7.40 (m, 5H), 5.04 (s, 2H), 4.10 (dd, $J$ = 13.2, 8.0 Hz, 1H), 3.63 (s, 3H), 2.31 (t, $J$ = 7.1 Hz, 2H), 1.91-2.03 (m, 1H), 1.78 (dt, $J$ = 14.5, 7.4 Hz, 1H); $^{13}C$ NMR (125 MHz, DMSO-$d_6$): δ= 173.6, 172.5, 156.1, 136.8, 128.3, 127.8, 127.7, 65.5, 53.0, 51.9, 29.8, 25.9; IR (neat): $\tilde{\upsilon}$ = 3513, 3345, 3032, 2956, 1711, 1529, 1454, 1438, 1258, 1210, 1176, 1047, 1024, 992, 818, 741, 698, 608, 580, 458 $cm^{-1}$; MS (ESI), *m/z*: calcd for $C_{14}H_{17}NNaO_6$, 318.1; found, 318.1 $[M + Na]^+$; $[\alpha]_D^{26}$ = - 12.4 (c = 1.09 in MeOH).

**1-Methyl-4-hydroxyethylcarbamoyl-N-Cbz-D-glutamic ester (37)**



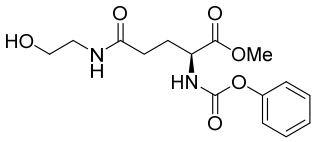
Following an identical procedure as the one described for compound **22**, starting from the mono-ester **36** (0.86 g, 2.71 mmol), alcohol **37** was obtained (0.54 g, 56% yield). $^{1}$H NMR (400 MHz, $CDCl_3$): δ= 7.19-7.27 (m, 5H), 6.72 (t, $J$ = 5.5 Hz, 1H), 6.13 (d, $J$ = 8.0 Hz, 1H), 4.99 (s, 2H), 4.23 (dt, $J$ = 8.3, 4.5 Hz, 1H), 3.61 (s, 3H), 3.49-3.57 (m, 2H), 3.15-3.32 (m, 2H), 2.04-2.21 (m, 3H), 1.87 (dt, $J$ = 15.3, 15.1, 6.8 Hz, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 172.7, 172.4, 156.4 136.0, 128.6, 128.3, 128.1, 67.2, 62.0, 53.1, 52.6, 42.5, 32.3, 28.8; IR (neat): $\tilde{\upsilon}$ = 3414, 3296, 3060, 3032, 2956, 1747, 1684, 1629, 1528, 1453, 1351, 1284, 1247, 1217, 1168, 1049, 864, 783, 756, 736, 697, 656, 613, 589, 428 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{16}H_{22}N_2NaO_6$, 361.1; found, 361.2 $[M + Na]^+$; $[\alpha]_D^{27}$ = - 15.4 (c = 1.09 in MeOH).

**1-Methyl-4-hydroxyethylcarbamoyl -D-glutamic ester (38)**



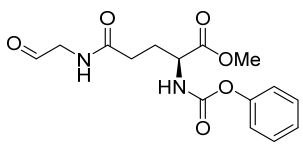
Following an identical procedure as the one described for compound **23**, starting from alcohol **36** (0.269 g, 0.76 mmol), amine **38** was obtained (0.137 g, 89 % yield). $^{1}$H NMR (500 MHz, MeOD): δ= 3.86 (s, 3H), 3.73 (t, $J$ = 5.7 Hz, 2H), 3.61 (t, $J$ = 6.9 Hz,1H), 3.43 (t, $J$ = 5.7 Hz, 2H), 2.45 (t, $J$ = 6.7 Hz, 2H), 2.12-2.21 (m, 1H), 1.96-2.07 (m, 1H); $^{13}$C NMR (100 MHz, MeOD): δ= 176.8, 175.3, 61.6, 54.7, 52.6, 43.0, 33.0, 31.5; IR (film): $\tilde{\upsilon}$ = 3362, 3281, 3078, 2933, 2864, 1732, 1644, 1547, 1437, 1368, 1205, 1176, 1060, 666, 574, 499 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_8H_{16}N_2$ $NaO_4$, 227.1; found, 227.0 $[M + Na]^+$;$[\alpha]_D^{27}$ = - 5.2 (c = 1.06 in MeOH).

**Phenyl-(S)-3-(2-hydroxyethylcarbamoyl)-1-(methoxycarbonyl)propylcarbamate (O6)**



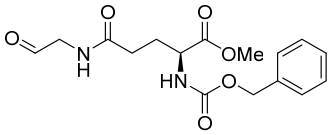
Following an identical procedure as the one described for compound **O3**, starting from amine **38** (49.8 mg, 0.244 mmol), carbamate **O6** was obtained (37.1 mg, 45 % yield). As a side product (11 %), the alcohol also reacted with phenyl cholorformate. $^{1}$H NMR (500 MHz, $CDCl_3$): δ= 7.35 (t, $J$ = 7.9 Hz, 2H), 7.20 (t, $J$ = 7.4 Hz, 1H), 7.12 (d, $J$ = 7.8 Hz, 2H), 6.46 (br, 1H), 6.26 (d, $J$ = 7.9 Hz, 1H), 4.40 (dt, $J$ = 8.4, 3.8 Hz, 1H), 3.77 (s, 3H), 3.62-3.71 (m, 2H), 3.41-3.47 (m, 1H), 3.27-3.33 (m, 1H), 2.66 (br, 1H), 2.26-2.38 (m, 3H), 1.98-2.05 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 172.9, 172.3, 154.8, 150.7, 129.3, 125.6, 121.5, 61.8, 53.5, 52.7, 42.5, 32.2, 28.4; IR (film): $\tilde{\upsilon}$ = 3330, 3286, 3055, 2945, 1720, 1649, 1535, 1488, 1437, 1351, 1260, 1201, 1140, 1068, 1025, 764, 756, 688, 496 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{15}H_{20}N_2$ $NaO_6$, 347.1; found, 347.1 $[M + Na]^+$; $[\alpha]_D^{28}$ = - 2.9 (c = 1.23 in DCM).

**Phenyl (S)-3-(formylmethylcarbamoyl)-1-(methoxycarbonyl)propylcarbamate (O4)**



Following an identical procedure as the one described for compound **O1**, starting from alcohol **O6** (13.0 mg, 0.038 mmol), aldehyde **O4** was obtained (6.3 mg, 49 % yield). $^{1}$H NMR (400 MHz, $CDCl_3$): δ= 9.64 (s, 1H), 7.36 (t, $J$ = 7.9 Hz, 2H), 7.20 (t, $J$ = 7.4 Hz, 1H), 7.13 (d, $J$ = 7.9 Hz, 2H), 6.54 (s, 1H), 6.02 (d, $J$ = 7.4 Hz, 1H), 4.49 (td, $J$ = 8.3, 4.3 Hz, 1H), 4.19 (qd, $J$ = 20.4, 5.1 Hz, 2H), 3.79 (s, 3H), 2.47-2.40 (m, 2H), 2.36-2.25 (m, 1H), 2.13-2.03 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 196.4, 172.3, 172.2, 154.7, 150.7, 129.3, 125.6 121.5 53.5 52.7, 50.3, 32.0, 28.6; IR (film): $\tilde{\upsilon}$ = 3333, 3055, 2956, 2927, 2852, 1723, 1652, 1530, 1488, 1455, 1438, 1345, 1202, 1141, 1025, 759, 689, 490, 403 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{15}H_{18}N_2NaO_6$, 345.1; found, 345.1 $[M + Na]^+$; $[\alpha]_D^{29}$ =- 4.9 (c = 0.27 in DCM).

**1-Methyl-4-formylmethylcarbamoyl-N-Cbz-L-glutamic ester (O5)**



Following an identical procedure as the one described for compound **O2**, starting from alcohol **37** (34.6 mg, 0.098 mmol), aldehyde **O5** was obtained (22.2 mg, 65 % yield). $^{1}$H NMR (500 MHz, $CDCl_3$): δ= 9.61 (s, 1H), 7.31-7.36 (m, 5H), 6.58 (br, 1H), 5.65 (d, $J$ = 7.8 Hz, 1H), 5.07-5.13 (m, 2H), 4.43 (dt, $J$ = 8.5, 4.5 Hz, 1H), 4.16 (dq, $J$ = 20.2, 4.9 Hz, 2H), 3.75 (s, 3H), 2.31-2.41 (m, 2H), 2.24 (td, $J$ = 19.2, 7.1 Hz, 1H), 1.94-2.01 (m, 1H); $^{13}$C NMR (125 MHz, $CDCl_3$): δ= 196.5, 172.7, 172.5, 173.3, 156.4, 136.0, 128.5, 128.3, 128.2, 128.1, 67.2, 67.1, 61.8, 53.3, 53.2, 52.6, 50.3, 42.5, 32.2, 32.0, 28.7, 28.7, 28.6, presence of rotamers; IR (film): $\tilde{\upsilon}$ = 3325, 3060, 3032, 2956, 1715, 1654, 1531, 1454, 1438, 1340, 1259, 1216, 1055, 740, 698, 592 cm$^{-1}$; MS (ESI), *m/z*: calcd for $C_{16}H_{20}N_2NaO_6$, 359.1; found, 359.1 $[M + Na]^+$; $[\alpha]_D^{29}$ = + 3.6 (c = 1.04 in DCM).

## 4.6.2 Cellular proliferation assays

Muscle NIH/3T3 (myc-), HT29 and MDA-MB-231 cells obtained from the UZH Cancer Institute and A498 cells obtained from Prof. Krek (Institute of Cell Biology, ETH) were cultured in DMEM supplemented with 10 % (v/v) fetal bovine serum, 100 units/mL of penicillin, 100 g/mL of streptomycin, 4.5 g/L glucose, 0.11g/L sodium pyrubate and 2mM glutamine and the cells were grown at 37 ºC in 5% $CO_2$ atmosphere with 80% relative humidity. A 5mM stock solution of the compounds in DMSO was prepared and kept at -20 °C.

The cells were plated at 5,000 cells per well (100 µL per well) in 96-well culture dishes and allowed to incubate for 24 h. The old media was removed, cells were washed with PBS (phosphate-buffered saline) and fresh media was added. A 5 mM solution of inhibitor (in

100% DMSO) was serially diluted in the culture media (12 different concentrations were used) and allowed to incubate for 24, 48 or 72 h. Control cells were treated with the same DMSO concentrations. After the incubation period the media was once more removed and the cells were washed with PBS to then be incubated with fresh media containing 86 nM resazurin. After 4 hours, the fluorescence was quantified using a fluorescence microplate reader (Biotek, FLx800$^{TM}$) at the respective excitation and emission wavelength of 560 and 590 nm. Resazurin becomes fluorescent under mitochondrial reduction and its emission directly correlates with the metabolic viability of the cells. The measured fluorescence values were corrected from the control samples containing DMSO and $IC_{50}$ values were determined by plotting the fraction of metabolically active cells against the log of the drug concentration. All measurements were performed at least in triplicate.

### 4.6.3 Fluorescent visualization of F-actin

10 000 cells per well in 90uL cell culture media were seeded in MEZEL diagnostic slides (PTFE coating around 10 wells of 6.7mm diameter each) and grown overnight. The cells were incubated with 4uM, 1uM and 250nM for 2 h or 8 h of the corresponding compound. For the recovery experiments, the cells were incubated for 2 h and allowed to recover in fresh media for 22 h.

The cells were washed with PBS and fixed with a solution of 4% paraformaldehyde in PBS for 10 minutes. Actin filaments were stained for 1h with a solution of 0.1 M tetramethyl rhodamine isothiocyanate TRITC labelled phalloidin (P1951, Sigma Aldrich) in PBS and cell nuclei were stained with a 1 g/mL DAPI (2-(4-amidinophenyl)-6-indolecarbamidine dihydrochloride) solution in PBS. The cells were mounted in Glycergel® mounting medium (Dako) after several PBS washes. Cells were visualized and photographed using a 40 x oil objective, A4 (nuclei) and TX2 (actin) filter cubes in a Leica wide field microscope (Hamamatsu EM-CCD).

### 4.6.4 *In vitro* actin polymerization experiments

In the polymerization experiments pyrene labeled muscle actin (Cytoeskeleton) was depolymerized on ice cold G-buffer (5 mM Tris-HCl pH 8.0, 0.2 mM $CaCl_2$ and 0.2 mM ATP) at a concentration of 2.5 µM for two days.

Actin was polymerized at 24 °C at 2.5 µM of the corresponding actin concentrations in a 96 well-plate (150 µL/well) and with 10 µM compound concentration. All the wells contained the same DMSO concentration (less than 5%). The fluorescence of pyrene labeled actin with the drug was monitored for 20 min to establish a baseline. Actin polymerization was induced by the addition of actin polymerization buffer (50 mM KCl, 2 mM $MgCl_2$ and 1 mM ATP, final concentrations). The polymerization process was

measured by monitoring the fluorescence once every 30 s for a total of 1 h-1.5 h in a Tecan Infinite® M1000 microplate reader with an excitation and emission wavelength of 350 nm and 407 nm respectively. Plotting and statistical analysis was performed using Origin.

## 4.7 References


1. Pollard, T. D.; Earnshaw, W. C.; Lippincott-Schwartz, J. *Cell Biology*. Saunders Elsevier: Philadelphia, 2007.
2. Image taken from Manual of Cellular and Molecular function. http://handbook.blueprint.org/.
3. Kingston, D. G. I. Taxol, a molecule for all seasons. *Chem. Commun.* **2001,** 10, 867-880.
4. Bray, D. *Cell movements*. Garland Publishing: New York & London, 1992.
5. Vandekerckhove, J.; Weber, K. At least six different actins are expressed in a higher mammal: an analysis based on the amino acid sequence of the amino-terminal tryptic peptide. *J. Mol. Biol.* **1978,** 126, 783-802.
6. Kabsch, W.; Mannherz, H. G.; Suck, D.; Pai, E. F.; Holmes, K. C. Atomic structure of the actin:DNase I complex. *Nature* **1990,** 347, 37-44.
7. Oriol, C.; Dubord, C.; Landon, F. Crystallization of native striated-muscle actin. *FEBS Lett.* **1977,** 73, 89-91.
8. Otterbein, L. R.; Graceffa, P.; Dominguez, R. The crystal structure of uncomplexed actin in the ADP state. *Science* **2001,** 293, 708-11.
9. Holmes, K. C.; Popp, D.; Gebhard, W.; Kabsch, W. Atomic model of the actin filament. *Nature* **1990,** 347, 44-9.
10. Oda, T.; Iwasa, M.; Aihara, T.; Maeda, Y.; Narita, A. The nature of the globular- to fibrous-actin transition. *Nature* **2009,** 457, 441-5.
11. Oosawa, F.; Asakura, S. *Thermodynamics of the polymerization of protein* Academic Press: London, New York, 1975.
12. Allingham, J. S.; Klenchin, V. A.; Rayment, I. Actin-targeting natural products: structures, properties and mechanisms of action. *Cell. Mol. Life Sci.* **2006,** 63, 2119-2134.
13. Pollard, T. D.; Mooseker, M. S. Direct measurement of actin polymerization rate constants by electron microscopy of actin filaments nucleated by isolated microvillus cores. *J. Cell Biol.* **1981,** 88, 654-9.
14. Higashi, S.; Oosawa, F. Conformational changes associated with polymerization and nucleotide binding in actin molecules. *J. Mol. Biol.* **1965,** 12, 843-65.
15. Wegner, A. Head to tail polymerization of actin. *J. Mol. Biol.* **1976,** 108, 139-50.
16. Kondo, H.; Ishiwata, S. Uni-directional growth of F-actin. *J. Biochem.* **1976,** 79, 159-71.
17. Carlier, M. F.; Pantaloni, D. Control of actin dynamics in cell motility. *J. Mol. Biol.* **1997,** 269, 459-67.

18. Cooper, J. A.; Pollard, T. D. Methods to measure actin polymerization. *Methods Enzymol.* **1982,** 85 Pt B, 182-210.
19. Carraway, K. L.; Carraway, C. A. C. *The Cytoeskeleton. A Practical Appoach.* . Oxford University Press: New York, 1992.
20. Lodish, H.; Baltimore, D.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Darnell, J. in *Mol. Cell Biol.*, 3rd Ed., Scientific American Books, New York. 1995.
21. Giganti, A.; Friederich, E. in *Prog. Cell Cycle Res.*, Vol. 5 (Eds.: L. Meijer, M. Roberge), Springer, Berlin. 2003.
22. Fenteany, G.; Zhu, S. Small-molecule inhibitors of actin dynamics and cell motility. *Curr. Top. Med. Chem.* **2003,** 3, 593-616.
23. Jordan, M. A.; Wilson, L. Microtubules and actin filaments: dynamic targets for cancer chemotherapy. *Curr. Opin. Cell Biol.* **1998,** 10, 123-30.
24. Brown, S. S. Cooperation between microtubule- and actin-based motor proteins. *Annu. Rev. Cell Dev. Biol.* **1999,** 15, 63-80.
25. Quinlan, R.; Hutchison, C.; Lane, B. *Intermediate Filament Proteins* (3rd Ed.), Oxford, UK: Oxford Univ. Press.
26. Wang, Y. L. Exchange of actin subunits at the leading edge of living fibroblasts: possible role of treadmilling. *J. Cell Biol.* **1985,** 101, 597-602.
27. Mitchison, T. J.; Cramer, L. P. Actin-based cell motility and cell locomotion. *Cell* **1996,** 84, 371-9.
28. Pollard, T. D.; Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **2003,** 112, 453-65.
29. Bretscher, M. S. Distribution of receptors for transferrin and low density lipoprotein on the surface of giant HeLa cells. *Proc. Natl. Acad. Sci. USA* **1983,** 80, 454-8.
30. Sellers, J. *Protein Profile* (2nd Ed.), edited by Sheterline P. Oxford, UK: Oxford Univ. Press. 1999.
31. Pollard, T. D.; Blanchoin, L.; Mullins, R. D. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Bioph. Biom.* **2000,** 29, 545-576.
32. Baum, J.; Papenfuss, A. T.; Baum, B.; Speed, T. P.; Cowman, A. F. Regulation of apicomplexan actin-based motility. *Nat. Rev. Microbiol.* **2006,** 4, 621-8.
33. Revenu, C.; Athman, R.; Robine, S.; Louvard, D. The co-workers of actin filaments: From cell structures to signals. *Nat. Rev. Mol. Cell Bio.* **2004,** 5, 635-646.
34. Higgs, H. N. Formin proteins: a domain-based approach. *Trends Biochem. Sci.* **2005,** 30, 342-53.
35. Quinlan, M. E.; Heuser, J. E.; Kerkhoff, E.; Mullins, R. D. Drosophila Spire is an actin nucleation factor. *Nature* **2005,** 433, 382-8.
36. Bamburg, J. R. Proteins of the ADF/cofilin family: essential regulators of actin dynamics. *Annu. Rev. Cell Dev. Biol.* **1999,** 15, 185-230.
37. Sun, H. Q.; Yamamoto, M.; Mejillano, M.; Yin, H. L. Gelsolin, a multifunctional actin regulatory protein. *J. Biol. Chem.* **1999,** 274, 33179-82.
38. Klenchin, V. A.; Allingham, J. S.; King, R.; Tanaka, J.; Marriott, G.; Rayment, I. Trisoxazole macrolide toxins mimic the binding of actin-capping proteins to actin. *Nat. Struct. Biol.* **2003,** 10, 1058-63.

39. Bubb, M. R.; Spector, I.; Bershadsky, A. D.; Korn, E. D. Swinholide A is a microfilament disrupting marine toxin that stabilizes actin dimers and severs actin filaments. *J. Biol. Chem.* **1995,** 270, 3463-6.
40. Terry, D. R.; Spector, I.; Higa, T.; Bubb, M. R. Misakinolide A is a marine macrolide that caps but does not sever filamentous actin. *J. Biol. Chem.* **1997,** 272, 7841-5.
41. Dauria, M. V.; Paloma, L. G.; Minale, L.; Zampella, A.; Verbist, J. F.; Roussakis, C.; Debitus, C.; Patissou, J. Reidispongiolide-a and Reidispongiolide-B, 2 New Potent Cytotoxic Macrolides from the New-Caledonian Sponge Reidispongia-Coerulea. *Tetrahedron* **1994,** 50, 4829-4834.
42. Zhang, X.; Minale, L.; Zampella, A.; Smith, C. D. Microfilament depletion and circumvention of multiple drug resistance by sphinxolides. *Cancer Res.* **1997,** 57, 3751-8.
43. Wada, S.; Matsunaga, S.; Saito, S.; Fusetani, N.; Watabe, S. Actin-binding specificity of marine macrolide toxins, mycalolide B and kabiramide D. *J. Biochem.* **1998,** 123, 946-52.
44. Saito, S.; Watabe, S.; Ozaki, H.; Kigoshi, H.; Yamada, K.; Fusetani, N.; Karaki, H. Novel actin depolymerizing macrolide aplyronine A. *J. Biochem.* **1996,** 120, 552-5.
45. Statsuk, A. V.; Bai, R.; Baryza, J. L.; Verma, V. A.; Hamel, E.; Wender, P. A.; Kozmin, S. A. Actin is the primary cellular receptor of bistramide A. *Nat. Chem. Biol.* **2005,** 1, 383-8.
46. Rizvi, S. A.; Courson, D. S.; Keller, V. A.; Rock, R. S.; Kozmin, S. A. The dual mode of action of bistramide A entails severing of filamentous actin and covalent protein modification. *Proc. Natl. Acad. Sci. USA* **2008,** 105, 4088-92.
47. Cooper, J. A. Effects of cytochalasin and phalloidin on actin. *J. Cell Biol.* **1987,** 105, 1473-8.
48. Tanaka, J.; Yan, Y.; Choi, J.; Bai, J.; Klenchin, V. A.; Rayment, I.; Marriott, G. Biomolecular mimicry in the actin cytoskeleton: mechanisms underlying the cytotoxicity of kabiramide C and related macrolides. *Proc. Natl. Acad. Sci. USA* **2003,** 100, 13851-6.
49. Spector, I.; Shochet, N. R.; Kashman, Y.; Groweiss, A. Latrunculins: novel marine toxins that disrupt microfilament organization in cultured cells. *Science* **1983,** 219, 493-5.
50. Coue, M.; Brenner, S. L.; Spector, I.; Korn, E. D. Inhibition of actin polymerization by latrunculin A. *FEBS Lett.* **1987,** 213, 316-8.
51. Morton, W. M.; Ayscough, K. R.; McLaughlin, P. J. Latrunculin alters the actin-monomer subunit interface to prevent polymerization. *Nat. Cell Biol.* **2000,** 2, 376-8.
52. Lorenz, M.; Popp, D.; Holmes, K. C. Refinement of the F-actin model against X-ray fiber diffraction data by the use of a directed mutation algorithm. *J. Mol. Biol.* **1993,** 234, 826-36.
53. Oda, T.; Namba, K.; Maeda, Y. Position and orientation of phalloidin in F-actin determined by X-ray fiber diffraction analysis. *Biophys. J.* **2005,** 88, 2727-36.
54. Steinmetz, M. O.; Stoffler, D.; Muller, S. A.; Jahn, W.; Wolpensinger, B.; Goldie, K. N.; Engel, A.; Faulstich, H.; Aebi, U. Evaluating atomic models of F-actin with an undecagold-tagged phalloidin derivative. *J. Mol. Biol.* **1998,** 276, 1-6.
55. Usui, T.; Kazami, S.; Dohmae, N.; Mashimo, Y.; Kondo, H.; Tsuda, M.; Terasaki, A. G.; Ohashi, K.; Kobayashi, J.; Osada, H. Amphidinolide h, a potent cytotoxic

macrolide, covalently binds on actin subdomain 4 and stabilizes actin filament. *Chem. Biol.* **2004,** 11, 1269-77.

56. Bubb, M. R.; Spector, I.; Beyer, B. B.; Fosen, K. M. Effects of jasplakinolide on the kinetics of actin polymerization. An explanation for certain in vivo observations. *J. Biol. Chem.* **2000,** 275, 5163-70.

57. Waldmann, H.; Hu, T. S.; Renner, S.; Menninger, S.; Tannert, R.; Oda, T.; Arndt, H. D. Total synthesis of chondramide C and its binding mode to F-actin. *Angew. Chem. Int. Ed.* **2008,** 47, 6473-7.

58. Bai, R.; Verdier-Pinard, P.; Gangwar, S.; Stessman, C. C.; McClure, K. J.; Sausville, E. A.; Pettit, G. R.; Bates, R. B.; Hamel, E. Dolastatin 11, a marine depsipeptide, arrests cells at cytokinesis and induces hyperpolymerization of purified actin. *Mol. Pharmacol.* **2001,** 59, 462-9.

59. Wehland, J.; Osborn, M.; Weber, K. Phalloidin-induced actin polymerization in the cytoplasm of cultured cells interferes with cell locomotion and growth. *Proc. Natl. Acad. Sci. USA* **1977,** 74, 5613-7.

60. Bai, R.; Covell, D. G.; Liu, C.; Ghosh, A. K.; Hamel, E. (-)-Doliculide, a new macrocyclic depsipeptide enhancer of actin assembly. *J. Biol. Chem.* **2002,** 277, 32165-71.

61. Saito, S. Y.; Feng, J.; Kira, A.; Kobayashi, J.; Ohizumi, Y. Amphidinolide H, a novel type of actin-stabilizing agent isolated from dinoflagellate. *Biochem. Biophys. Res. Commun.* **2004,** 320, 961-5.

62. Marquez, B. L.; Watts, K. S.; Yokochi, A.; Roberts, M. A.; Verdier-Pinard, P.; Jimenez, J. I.; Hamel, E.; Scheuer, P. J.; Gerwick, W. H. Structure and absolute stereochemistry of hectochlorin, a potent stimulator of actin assembly. *J. Nat. Prod.* **2002,** 65, 866-71.

63. Oda, T.; Crane, Z. D.; Dicus, C. W.; Sufi, B. A.; Bates, R. B. Dolastatin 11 connects two long-pitch strands in F-actin to stabilize microfilaments. *J. Mol. Biol.* **2003,** 328, 319-24.

64. Spector, I.; Braet, F.; Shochet, N. R.; Bubb, M. R. New anti-actin drugs in the study of the organization and function of the actin cytoskeleton. *Microsc. Res. Tech.* **1999,** 47, 18-37.

65. Hayot, C.; Debeir, O.; Van Ham, P.; Van Damme, M.; Kiss, R.; Decaestecker, C. Characterization of the activities of actin-affecting drugs on tumor cell migration. *Toxicol. Appl. Pharmacol.* **2006,** 211, 30-40.

66. Shi, Q.; Chen, K.; Morris-Natschke, S. L.; Lee, K. H. Recent progress in the development of tubulin inhibitors as antimitotic antitumor agents. *Curr. Pharm. Des.* **1998,** 4, 219-48.

67. Allingham, J. S.; Zampella, A.; D'Auria, M. V.; Rayment, I. Structures of microfilament destabilizing toxins bound to actin provide insight into toxin design and activity. *Proc. Natl. Acad. Sci. USA* **2005,** 102, 14527-32.

68. Newman, D. J.; Cragg, G. M.; Snader, K. M. Natural products as sources of new drugs over the period 1981-2002. *J. Nat. Prod.* **2003,** 66, 1022-37.

69. Koch, M. A.; Waldmann, H. Protein structure similarity clustering and natural product structure as guiding principles in drug discovery. *Drug Discov. Today* **2005,** 10, 471-83.

70. http://zinc.docking.org/.
71. Maass, P.; Schulz-Gasch, T.; Stahl, M.; Rarey, M. Recore: a fast and versatile method for scaffold hopping based on small molecule crystal structure conformations. *J. Chem. Inf. Model* **2007,** 47, 390-9.
72. Oloff, S.; Zhang, S.; Sukumar, N.; Breneman, C.; Tropsha, A. Chemometric analysis of ligand receptor complementarity: identifying Complementary Ligands Based on Receptor Information (CoLiBRI). *J. Chem. Inf. Model* **2006,** 46, 844-51.
73. Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed.* **2002,** 41, 2644-76.
74. Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999,** 42, 5100-5109.
75. Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. Consensus scoring for ligand/protein interactions. *J. Mol. Graph. Model* **2002,** 20, 281-295.
76. Clark, M.; Cramer, R. D.; Vanopdenbosch, N. Validation of the General-Purpose Tripos 5.2 Force-Field. *J. Comput. Chem.* **1989,** 10, 982-1012.
77. Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhardt, C.; Caflisch, A. Exhaustive docking of molecular fragments with electrostatic solvation. *Tetrahedron Lett.* **1999,** 37, 88-105.
78. Majeux, N.; Scarsi, M.; Caflisch, A. Efficient electrostatic solvation model for protein-fragment docking. *Proteins* **2001,** 42, 256-268.
79. Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009,** 30, 1545-614.
80. Im, W.; Beglov, D.; Roux, B. Continuum Solvation Model: computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation. *Comput. Phys. Commun.* **1998,** 111, 59-75.
81. Klon, A. E.; Glick, M.; Davies, J. W. Combination of a naive Bayes classifier with consensus scoring improves enrichment of high-throughput docking results. *J. Med. Chem.* **2004,** 47, 4356-4359.
82. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliver. Rev.* **1997,** 23, 3-25.
83. Trott, O.; Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010,** 31, 455-461.
84. Cooper, J. A.; Walker, S. B.; Pollard, T. D. Pyrene actin: documentation of the validity of a sensitive assay for actin polymerization. *J. Muscle Res. Cell Motil.* **1983,** 4, 253-62.

85. Furstner, A.; Kirk, D.; Fenster, M. D. B.; Aissa, C.; De Souza, D.; Muller, O. Diverted total synthesis: Preparation of a focused library of latrunculin analogues and evaluation of their actin-binding properties. *Proc. Natl. Acad. Sci. USA* **2005,** 102, 8103-8108.

86. Hidaka, K.; Kimura, T.; Ruben, A. J.; Uemura, T.; Kamiya, M.; Kiso, A.; Okamoto, T.; Tsuchiya, Y.; Hayashi, Y.; Freire, E.; Kiso, Y. Antimalarial activity enhancement in hydroxymethylcarbonyl (HMC) isostere-based dipeptidomimetics targeting malarial aspartic protease plasmepsin. *Bioorgan. Med. Chem.* **2008,** 16, 10049-10060.

87. Chen, H. L.; Feng, Y. Q.; Xu, Z. S.; Xu, Z. S.; Ye, T. The total syntheses and reassignment of stereochemistry of dragonamide. *Tetrahedron* **2005,** 61, 11132-11140.

88. Still, W. C.; Gennari, C. Direct Synthesis of Z-Unsaturated Esters - a Useful Modification of the Horner-Emmons Olefination. *Tetrahedron Lett.* **1983,** 24, 4405-4408.

89. Tu, G. G.; Li, S. H.; Huang, H. M.; Li, G.; Xiong, F.; Mai, X.; Zhu, H. W.; Kuang, B. H.; Xu, W. F. Novel aminopeptidase N inhibitors derived from 1,3,4-thiadiazole scaffold. *Bioorgan. Med. Chem.* **2008,** 16, 6663-6668.

90. Marecek, J.; Song, B.; Brewer, S.; Belyea, J.; Dyer, R. B.; Raleigh, D. P. A simple and economical method for the production of C-13, O-18-labeled Fmoc-amino acids with high levels of enrichment: Applications to isotope-edited IR studies of proteins. *Org. Lett.* **2007,** 9, 4935-4937.

91. Tantry, S. J.; Venkataramanarao, R.; Chennakrishnareddy, G.; Sureshbabu, V. V. Total synthesis of cyclosporin O by convergent approach employing Fmoc-Amino acid chlorides mediated by zinc dust. *J. Org. Chem.* **2007,** 72, 9360-9363.

92. Kalesh, K. A.; Tan, L. P.; Lu, K.; Gao, L. Q.; Wang, J. G.; Yao, S. Q. Peptide-based activity-based probes (ABPs) for target-specific profiling of protein tyrosine phosphatases (PTPs). *Chem. Commun.* **2010,** 46, 589-591.

93. Percec, V.; Dulcey, A. E.; Balagurusamy, V. S. K.; Miura, Y.; Smidrkal, J.; Peterca, M.; Nummelin, S.; Edlund, U.; Hudson, S. D.; Heiney, P. A.; Hu, D. A.; Magonov, S. N.; Vinogradov, S. A. Self-assembly of amphiphilic dendritic dipeptides into helical pores. *Nature* **2004,** 430, 764-768.

94. Mandal, A. B.; Jayakumar, R. A New Micelle-Forming Peptide. *J. Chem. Soc. Chem. Comm.* **1993**, 237-238.

95. Yang, H. Z.; Xu, S.; Liao, X. Y.; Zhang, S. D.; Liang, Z. L.; Liu, B. H.; Bai, J. Y.; Jiang, C.; Ding, J.; Cheng, G. F.; Liu, G. A novel immunostimulator, N-2-[(alpha-O-benzyl-N-(acetylmuramyl)-L-alanyl-D-isoglutaminyl]-N-6-trans-(m-nitrocinnamoyl)-L-lysine, and its adjuvancy on the hepatitis B surface antigen. *J. Med. Chem.* **2005,** 48, 5112-5122.

96. Luesch, H.; Hoffmann, D.; Hevel, J. M.; Becker, J. E.; Golakoti, T.; Moore, R. E. Biosynthesis of 4-methylproline in cyanobacteria: Cloning of nosE and nosF genes and biochemical characterization of the encoded dehydrogenase and reductase activities. *J. Org. Chem.* **2003,** 68, 83-91.

97. Hanessian, S.; Sahoo, S. P. A Novel and Efficient Synthesis of L-Vinylglycine. *Tetrahedron Lett.* **1984,** 25, 1425-1428.

# 5

# CONCLUSIONS AND OUTLOOK

Modern drug discovery involves the identification of hits for a particular target, medicinal chemistry optimization of those hits to increase the affinity, selectivity, efficacy and drug-like properties, followed by *in vivo* efficacy studies preceding clinical development. In this Ph.D. thesis, three independent projects dealing with *in silico* fragment-based hit discovery and medicinal chemistry optimization campaigns are presented for the discovery of kinase, bromodomain and actin inhibitors.

First, single digit nanomolar potent kinase inhibitors were designed based on X-ray structures of EphA3 kinase in complex with *in silico* identified hits. Different binding modes were designed, which resulted in the development of a total of 25 type I, $I_{1/2}$ and II kinase inhibitors. These inhibitors were characterized by an array of biophysical techniques including enzymatic assays, differential scanning fluorimetry and surface plasmon resonance. Some of these compounds presented good selectivity profiles against a panel of 453 kinases and their antiproliferation activity against a panel of cancer cells revealed remarkable growth inhibition values in the low nanomolar range, especially against a leukemia cell line. Finally, *in vivo* assays (mice xenografted with human breast cancer) confirmed the cytostatic activity of one of our inhibitors, which makes it a promising candidate lead for further preclinical development. Further steps to be carried out include the evaluation of solubility and metabolic stability of the designed inhibitors, as well as their study on a leukemia mouse model due to the high antiproliferative activities observed in a leukemia cell line for several derivatives.

The next project presented in this thesis aimed to identify and optimize CREBBP bromodomain ligands. Bromodomains have recently been recognized as potential therapeutic targets, and thus, there is a big need for the development of potent and selective bromodomain ligands that could validate this approach. During this thesis, novel ligands of the CREBBP bromodomain were identified by fragment-based docking. The *in silico* hits were optimized by chemical synthesis resulting in 24 derivatives, some of which

turned out to be selective nanomolar ligands achieving an almost 30 fold improvement in binding affinity. The synthesized analogues presented unprecedented selectivity values for the CREBBP bromodomain over other human bromodomain sub-families, as a result of establishing electrostatic interactions between the designed ligands and the Arg1173 side chain of the protein. The mentioned interaction was confirmed by solving the high-resolution crystal structure of the CREBBP bromodomain in complex with one of the synthesized ligands. A preliminary cell based screening revealed that our CREBBP ligands are able to inhibit the proliferation of leukemia cell lines in the low micromolar level. Future research will concentrate on the study of properties such as water solubility, by the addition of water solubilizing groups (e.g. morpholine, piperazine, etc.) and membrane permeability, by the substitution of the benzoic acid moiety present in the compounds by isosters (e. g. tetrazole, sulfonamides etc.) in order to develop promising lead candidates whose efficacy will be evaluated on *in vivo* models.

Another aspect of this Ph. D. work dealt with the *in silico* design, synthesis and biological evaluation of actin binders. Actin is a novel and attractive protein target that has so far remained unexplored in the context of small organic molecule ligands. During this work, we have designed and validated a new computational approach to discover novel actin leads targeting the ATP binding site of actin. The *in silico* campaign resulted in the selection of fifteen fragments, from which five scaffolds were chemically synthesized. Cell based studies revealed that the compounds were able to alter the morphology of the actin cytoskeleton at low micromolar concentrations by reducing the microfilament bundles (stress fibers) and inducing cell shrinkage and actin aggregation. Moreover, moderate inhibition of actin polymerization was obtained in an *in vitro* actin polymerization assay. As a result, the developed small organic molecules are promising starting hits for the development of more potent actin inhibitors. In the future we aim to design and synthesize more potent derivatives of the mentioned scaffolds, as well as to implement a computer based *de novo* design campaign aiming to develop novel actin binding molecules.

To sum up, we can conclude that the combination of computer-based drug design, X-ray crystallography, chemical synthesis and *in vitro* and cell based assays performed during this work has proven to be an efficient strategy for the development of promising kinase and bromodomain ligands as well as for the identification of actin inhibitor hits with moderate potency.

# CURRICULUM VITAE

Surname: UNZUE LOPEZ

Name: ANDREA

Date of birth: January 28th, 1987

Nationality: Spanish

## Education

| | |
|---|---|
| 2010 – 2015 | **Fast track PhD in Organic Chemistry, University of Zürich, Switzerland**<br>''Small Organic Molecules as Tunable Tools for Biology''. Prof. Nevado's group. |
| 2010 – 2012 | **Master of Science in Chemical and Molecular Sciences, University of Zurich, Zurich, Switzerland**<br>''Iriomoteolides as Novel Chemical Tools for the Study of Actin Dynamics''. Prof. Nevado's group. |
| 2009 – 2010 | **Diploma thesis, University of Wisconsin-Madison, USA**<br>''New Methods for Polyamine Synthesis''. Prof. Schomaker's group. |
| 2005 – 2010 | **Diploma in Chemistry, University of the Basque Country University, San Sebastian, Spain**<br>Overall grade of 9.28 out of 10. |

## Publications

[I] A. Unzue, J. Dong, K. Lafleur, H. Zhao, E. Frugier, A. Caflisch, and C. Nevado, ''Pyrrolo[3,2-b]quinoxaline Derivatives as Types I1/2 and II Eph Tyrosine Kinase Inhibitors: Structure-Based Design, Synthesis, and *in Vivo* Validation,'' *Journal of Medicinal Chemistry*, 2014, Vol. 57, No. 15, pp. 6834–6844.

[II] A. Unzue, K. Lafleur, H. Zhao, T. Zhou, J. Dong, P. Kolb, J. Liebl, S. Zahler, A. Caflisch, and C. Nevado, ''Three stories on Eph kinase inhibitors: from *in silico* discovery to *in vivo* validation,'' *Submitted to Current Drug Targets in November 2014.*

[III] A. Unzue, J. Dong, K. Lafleur, H. Zhao, E. Frugier, A. Caflisch, and C. Nevado, ''2-amino-1-phenyl-pyrrolo[3,2-b]quinoxaline-3-carboxamide derivatives,'' *Patent application number 14162145.8. Patent number 1462. Patent pending.*

[IV] A. Unzue, M. Xu, J. Dong, L. Wiedmer, D. Spiliotopoulos, A. Caflisch, and C. Nevado, ''Fragment-based Design of Selective Nanomolar Ligands of the CREBBP Bromodomain,'' *Submitted to Angewandte Chemie in December 2014.*

[V] M. Xu, A. Unzue, J. Dong, D. Spiliotopoulos, C. Nevado, and A. Caflisch, ''CREBBP bromodomain ligands discovered by docking, optimized by molecular dynamics-guided synthesis of derivatives, and validated by X-ray crystallography,'' *Manuscript in preparation.*

[VI] A. Unzue, M. Xu, J. Dong, D. Spiliotopoulos, A. Caflisch, and C. Nevado, ''Compounds, in particular for use in the treatment of a disease or condition for which a bromodomain inhibitor is indicated,'' *Patent pending.*

[VII] A. Anson-Casaos, I. Tacchini, A. Unzue, and M. T. Martinez, ''Combined modification of a $TiO_2$ photocatalyst with two different carbon forms,'' *Applied Surface Science*, 2013, Vol. 270, pp. 675-684.

## Poster Presentation/Conferences

Oct. 2014 **9th Dorothy Crowdfoot Hodgkin Interdisciplinary Symposium (UZH-Zürich)**
Poster: ''Pyrrolo[3,2-b]quinoxaline Derivatives as Types I1/2 and II Eph Tyrosine Kinase Inhibitors''
A. Unzue, J. Dong, K. Lafleur, H. Zhao, E. Frugier, A. Caflisch, C. Nevado.

Sept. 2014 **1st EFMC-YOUNG Medicinal Chemist Symposium (Lisbon, Portugal)**
Oral presentation: ''In silico Designed CREBBP Bromodomain Selective Inhibitors''
A. Unzue, M. Xu, J. Dong, D. Spiliotopoulos, A. Caflisch, C. Nevado.

Sept. 2014 **XXIII International Symposium on Medicinal Chemistry (Lisbon, Portugal)**
Poster: ''In silico Designed CREBBP Bromodomain Selective Inhibitors''
A. Unzue, M. Xu, J. Dong, D. Spiliotopoulos, A. Caflisch, C. Nevado.

Oct. 2013 **8th Dorothy Crowdfoot Hodgkin Interdisciplinary Symposium (UZH-Zürich)**
Poster: ''Iriomoteolides as Chemical Tools for the Study of Actin Dynamics'' A. Unzue, R. Cribiu, M. Maluszynska-Hoffman, C. Nevado.

Aug. 2013 **Swiss Summer School 2013 in Synthesis and Catalysis (Villars sur Ollon, Switzerland)**
Oral presentation: ''Iriomoteolides as Chemical Tools for the Study of Actin Dynamics''
A. Unzue, R. Cribiu, M. Maluszynska-Hoffman, C. Nevado.

Sept. 2012 **Swiss Chemical Society Meeting (ETH-Zürich)**
Poster: ''*In silico* based actin inhibitors''
A. Unzue, S. Rennebaum, A. Caflisch, C. Nevado.

May 2010 **American Chemical Society Meeting (San Francisco, USA)**
Poster: ''New Methods for Polyamine Synthesis''
A. Unzue, D. Marston, C. Fry, J. Sun, J. Rigoli, J. M. Schomaker.

## Awards

- Poster award at the 9th Dorothy Crowfoot Hodgkin Interdisciplinary Symposium, Zürich, Oct. 2014.
- 2014 SCNAT/SCS Chemistry Travel Award (May 2014).
- Poster award at the 8th Dorothy Crowfoot Hodgkin Interdisciplinary Symposium, Zürich, Oct. 2013.
- Prof. Dr. Hans E. Schmid-Stiftung grant from the UZH during the Master in Chemical and Molecular Sciences.
- Best chemistry-diploma award from the University of the Basque Country (founded by Kutxa) in Chemistry. November 2010.
- JAE-INTRO grant from The Spanish National Research Council (CSIC). July-August 2009.